Biotech-Chem Library

Questions about the scope or the results of the search? Contact **the searcher or contact**:

Mary Hale, Information Branch Supervisor
308-4258, CM1-1E01

Voluntary Results Feedback Form

➢ *I am an examiner in Workgroup:* [   ]   *Example: 1610*

➢ *Relevant prior art **found**, search results used as follows:*

☐ 102 rejection

☐ 103 rejection

☐ Cited as being of interest.

☐ Helped examiner better understand the invention.

☐ Helped examiner better understand the state of the art in their technology.

   Types of relevant prior art found:

   ☐ Foreign Patent(s)

   ☐ Non-Patent Literature
      (journal articles, conference proceedings, new product announcements etc.)

➢ *Relevant prior art **not found**:*

☐ Results verified the lack of relevant prior art (helped determine patentability).

☐ Results were not useful in determining patentability or understanding the invention.

Comments:

sira
Search and Information
Resources Administration

From: Slobodyansky, Elizabeth
Sent: Tuesday, September 23, 2003 6:42 PM
To: STIC-Biotech/ChemLib
Subject: 09/974,973

Please search for case 09/974,973

SEQ ID NOs: 2 and 19 against commercial ~~and interference~~ databases.

Thank you.

Elizabeth Slobodyansky, PhD

Primary Examiner

Art Unit 1652
CM1 10D11
703-306-3222

mail box 10D01

Searcher:_____	TYPE OF SEARCH:	VENDOR/COST (where applic.)
Phone:_____	NA Sequences:_____	STN:_____
Location:_____	AA Sequences:_____	DIALOG:_____
Date Picked Up:_____	Structures:_____	Questel/Orbit:_____
Date Completed:_____	Bibliographic:_____	DRLink:_____
Searcher Prep/Review:_____	Litigation:_____	Lexis/Nexis:_____
Clerical:_____	Full text:_____	Sequence Sys.:_____
Online time:_____	Patent Family:_____	WWW/Internet:_____
	Other:_____	Other (specify):_____

OM protein - protein search, using sw model

```
        GenCore version 5.1.6
        Copyright (c) 1993 - 2003 Compugen Ltd.
```

Run on: September 24, 2003, 15:43:18 ; Search time 64.4737 Seconds
                        (without alignments)
                        2848.398 Million cell updates/sec

Title: US-09-974-973A-2
Perfect score: 5865
Sequence: 1 MTAITLGGLLLKGIITLVST........RVVVPAATKVEGGDLIVVVS 1157

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   A_Geneseq_19Jun03:*

```
        1: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1980.DAT:*
        2: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1981.DAT:*
        3: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1982.DAT:*
        4: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1983.DAT:*
        5: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1984.DAT:*
        6: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1985.DAT:*
        7: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1986.DAT:*
        8: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1987.DAT:*
        9: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1988.DAT:*
       10: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1989.DAT:*
       11: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1990.DAT:*
       12: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1991.DAT:*
       13: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1992.DAT:*
       14: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1993.DAT:*
       15: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1994.DAT:*
       16: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1995.DAT:*
       17: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1996.DAT:*
       18: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1997.DAT:*
       19: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1998.DAT:*
       20: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1999.DAT:*
       21: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2000.DAT:*
       22: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2001.DAT:*
       23: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2002.DAT:*
       24: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	
1	5865	100.0	1157	23	AAU98050	Corynebacterium fe	
2	5779	98.5	1140	23	AAU98052	Corynebacterium mu	
3	5764	98.3	1141	21	AAB01436	Pyruvate carboxyla	
4	5759	98.2	1140	22	AAG90511	C glutamicum prote	
5	5759	98.2	1140	23	AAB67129	Corynebacterium gl	
6	5759	98.2	1140	23	AAE25601	Corynebacterium gl	
7	5759	98.2	1140	23	AAU98053	Corynebacterium wi	
8	5753	98.1	1140	20	AAW93971	C. glutamicum pyru	
9	5751	98.1	1140	22	AAG93249	C glutamicum prote	
10	5299.5	90.4	1139	22	AAB83180	Corynebacterium th	
11	2678	45.7	532	22	AAB79302	Corynebacterium gl	
12	2530	43.1	1148	22	AAU00511	Bacillus subtilis	
13	2494	42.5	1146	23	ABB47612	Listeria monocytog	
14	2489	42.4	1147	22	AAG33972	Staphylococcus aur	
15	2465	42.0	1151	23	AAG82677	S. epidermidis ope	
16	2465	42.0	1154	23	ABP38583	Staphylococcus epi	
17	2460.5	42.0	1142	22	AAG35213	Enterococcus faeca	
18	2441	41.6	1181	22	ABB58211	Drosophila melanog	
19	2441	41.6	1181	22	ABB66604	Drosophila melanog	
20	2441	41.6	1181	22	ABB66605	Drosophila melanog	
21	2433.5	41.5	1196	22	ABB67309	Drosophila melanog	
22	2412	41.1	1136	23	ABB53980	Lactococcus lactis	
23	2336	39.8	461	22	AAB79303	Corynebacterium gl	
24	2300	39.2	1073	22	AAU36768	Staphylococcus aur	
25	1546.5	26.4	320	22	AAB79300	Corynebacterium gl	
26	1546.5	26.4	320	22	AAB79301	Corynebacterium gl	
27	1371	23.4	272	22	AAB79298	Corynebacterium gl	
28	1371	23.4	272	22	AAB79299	Corynebacterium gl	
29	1032.5	17.6	447	17	AAW05207	Anabaena biotin bi	
30	1032.5	17.6	447	17	AAW70402	Anabaena biotin ca	
31	1032.5	17.6	448	15	AAR51080	Biotin carboxylase	
32	1005.5	17.1	453	15	AAR51083	fabG gene encoding	
33	1005.5	17.1	453	17	AAW05208	Synechococcus biot	
34	1005.5	17.1	453	19	AAW70403	Synechococcus biot	
35	980	16.7	425	21	AAB11663	A. vitis hypersens	
36	979.5	16.7	471	22	AAU33719	Pseudomonas aerugi	
37	956	16.3	449	22	AAU38292	Salmonella typhi c	
38	954.5	16.3	455	23	ABB54091	Lactococcus lactis	
39	950	16.2	448	22	AAU35567	Haemophilus influe	
40	947.5	16.2	456	23	ABP28017	Streptococcus poly	
41	940	16.0	449	16	AAR66742	Biotin-carboxylase	
42	940	16.0	449	22	AAU34737	E. coli cellular p	
43	938.5	16.0	453	24	ABP77556	N. gonorrhoeae ami	
44	937.5	16.0	454	20	AAY34779	Chlamydia pneumoni	
45	936.5	16.0	455	22	AAU38024	Streptococcus pneu	

ALIGNMENTS

```
RESULT 1
AAU98050
ID   AAU98050 standard; Protein; 1157 AA.
XX
AC   AAU98050;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Corynebacterium feedback-resistant pyruvate carboxylase enzyme.
XX
KW   Feedback-resistant; pyruvate carboxylase; enzyme;
KW   aspartic acid feedback inhibition resistant.
XX
OS   Corynebacterium glutamicum.
XX
FH   Key             Location/Qualifiers
FT   Region          1..18
FT                   /note="Specifically claimed in claim 18"
FT   Region          164..176
FT                   /note="Specifically claimed in claim 18"
FT   Region          193..205
FT                   /note="Specifically claimed in claim 18"
FT   Region          217..229
FT                   /note="Specifically claimed in claim 18"
FT   Region          238..250
FT                   /note="Specifically claimed in claim 18"
FT   Region          466..478
FT                   /note="Specifically claimed in claim 18"
XX
PN   WO200231158-A2.
XX
PD   18-APR-2002.
```

```
XX
PF   12-OCT-2001; 2001WO-US31893.
PR   13-OCT-2000; 2000US-239913P.
XX
PA   (ARCH ) ARCHER-DANIELS MIDLAND CO.
XX
PI   Hanke PD;
XX
DR   WPI; 2002-463267/49.
DR   N-PSDB; ABK52832.
XX
PT   Novel mutated, feedback resistant pyruvate carboxylase enzyme
PT   polypeptide, useful for producing amino acids e.g. L-lysine,
PT   L-threonine, L-glycine, L-glutamic acid, L-proline and L-methionine and
PT   L-isoleucine.
XX
PS   Claim 17; Fig 1; 42pp; English.
XX
CC   The present invention relates to a new mutated, feedback-resistant
CC   pyruvate carboxylase enzyme. The invention is useful for producing an
CC   amino acid (e.g. L-Lys, L-Thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro),
CC   by culturing a host cell in a suitable media and separating the amino
CC   acid from the medium. The vector of the invention is useful for
CC   replacement of a wild-type pyruvate carboxylase gene, with a feedback
CC   resistant pyruvate carboxylase gene, in a Corynebacterium glutamicum, by
CC   replacing a genomic copy of the wild-type pyruvate carboxylase gene with
CC   a selectable marker gene through homologous recombination to form a first
CC   recombination strain, and replacing the selectable marker gene in the
CC   first recombinant strain, with feedback resistant pyruvate carboxylase
CC   gene through homologous recombination to form a second recombinant
CC   strain, where the homologous recombination in the above steps, occurs
CC   between the host cell and the vector. The feedback-resistant pyruvate
CC   carboxylase enzyme is resistant to feedback inhibition from aspartic
CC   acid. The present amino acid sequence represents the feedback-resistant
CC   pyruvate carboxylase enzyme of the invention.
XX
SQ   Sequence   1157 AA;

Query Match      100.0%;   Score 5865;   DB 23;   Length 1157;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1157;   Conservative    0;   Mismatches    0;   Indels    0;   Gaps    0;

Qy     1  MTAITLGGLLKGIITLVSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MTAITLGGLLKGIITLVSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYP 60

Qy    61  REDRGSFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  REDRGSFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLA 120

Qy   121  RECAENGITFIGPTPEVLLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTY 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  RECAENGITFIGPTPEVLLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTY 180

Qy   181  PIFYKAVAGGGGRGMRFVSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQI 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  PIFYKAVAGGGGRGMRFVSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQI 240

Qy   241  LGDRTGEVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  LGDRTGEVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGT 300

Qy   301  VEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  VEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKT 360

Qy   361  HGAALQCRITTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  HGAALQCRITTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMT 420

Qy   421  CRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  CRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQA 480

Qy   481  PPADDEQGRILDYLADVTVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  PPADDEQGRILDYLADVTVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAA 540

Qy   541  FARDLREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATY 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  FARDLREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATY 600

Qy   601  DVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVD 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  DVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVD 660

Qy   661  IFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIV 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  IFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIV 720

Qy   721  KSGAHILAIKDMAGLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGAD 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  KSGAHILAIKDMAGLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGAD 780

Qy   781  AVDGASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  AVDGASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGT 840

Qy   841  PGPTGRVYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVV 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  PGPTGRVYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVV 900

Qy   901  GDLALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  GDLALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEG 960

Qy   961  KAPLTEVPEEQAHLDADDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  KAPLTEVPEEQAHLDADDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYG 1020

Qy  1021  LVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTAT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  LVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTAT 1080

Qy  1081  AEKADSSNKGHVAAPFAGVVTVTVAEGDEVKAGDAVIIEAMKMEATITASVDGKIERVV 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  AEKADSSNKGHVAAPFAGVVTVTVAEGDEVKAGDAVIIEAMKMEATITASVDGKIERVV 1140

Qy  1141  VPAATKVEGGDLIVVVS 1157
          |||||||||||||||||
Db  1141  VPAATKVEGGDLIVVVS 1157


RESULT 2
AAU98052
ID   AAU98052 standard; Protein; 1140 AA.
XX
AC   AAU98052;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Corynebacterium mutant feedback-resistant pyruvate carboxylase enzyme.
XX
KW   Feedback-resistant; pyruvate carboxylase; enzyme;
KW   aspartic acid feedback inhibition resistant; mutant; mutein.
XX
OS   Corynebacterium glutamicum.
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "Wild-type Met substituted by Val"
FT   Misc-difference 153
FT   Misc-difference 182
FT                   /note= "Wild-type Glu substituted by Asp"
```

```
FT   Misc-difference   206
FT                     /note= "Wild-type Ala substituted by Ser"
FT   Misc-difference   227
FT                     /note= "Wild-type Ala substituted by Ser"
FT   Misc-difference   455
FT                     /note= "Wild-type His substituted by Arg"
FT   Region            /note= "Wild-type Ala substituted by Gly"
FT                     1110..1122
FT   Misc-difference   1116
FT                     /note= "Specifically claimed in claim 18"
FT                     /note= "Wild-type Asp substituted by Glu"
XX
PN   WO200231158-A2.
XX
PD   18-APR-2002.
XX
PF   12-OCT-2001; 2001WO-US31893.
XX
PR   13-OCT-2000; 2000US-239913P.
XX
PA   (ARCH ) ARCHER-DANIELS MIDLAND CO.
XX
PI   Hanke PD;
XX
DR   WPI; 2002-463267/49.
XX
PT   Novel mutated, feedback resistant pyruvate carboxylase enzyme
PT   polypeptide, useful for producing amino acids e.g. L-lysine,
PT   L-threonine, L-glycine, L-glutamic acid, L-proline and L-methionine and
PT   L-isoleucine -
XX
PS   Claim 1; Page -; 42pp; English.
XX
CC   The present invention relates to a new mutated, feedback-resistant
CC   Pyruvate carboxylase enzyme. The invention is useful for producing an
CC   amino acid (e.g. L-Lys, L-Thr, L-Ile, L-Glu, L-Arg and L-Pro),
CC   by culturing a host cell in a suitable media and separating the amino
CC   acid from the medium. The vector of the invention is useful for
CC   replacement of a wild-type pyruvate carboxylase gene, with a feedback
CC   resistant pyruvate carboxylase gene, in a Corynebacterium glutamicum, by
CC   replacing a genomic copy of the wild-type pyruvate carboxylase gene with
CC   a selectable marker gene through homologous recombination to form a first
CC   recombination strain, and replacing the selectable marker gene in the
CC   first recombinant strain, with feedback resistant pyruvate carboxylase
CC   gene through homologous recombination to form a second recombinant
CC   strain, where the homologous recombination in the above steps, occurs
CC   between the host cell and the vector. The feedback-resistant pyruvate
CC   carboxylase enzyme is resistant to feedback inhibition from aspartic
CC   acid. The present amino acid sequence represents the mutant
CC   feedback-resistant pyruvate carboxylase enzyme of the invention.
CC   Note: The present sequence is not shown in the specification but is
CC   derived from the wild-type feedback-resistant pyruvate carboxylase
CC   enzyme (AAU98053) given in figure 2 of the specification.
XX
SQ   Sequence  1140 AA;

Query Match             98.5%;  Score 5779;  DB 23;  Length 1140;
Best Local Similarity   99.8%;
Matches 1138;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy   18 VSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 77
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 VSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60

Qy   78 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 137
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120

Qy  138 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRF 197
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRF 180

Qy  198 VSSPDELRKLATEASREEAAAFGDGSVIVERAVINPQHIEVQILGDRTGEVHLYERDCS 257
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 VSSPDELRKLATEASREEAAAFGDGSVIVERAVINPQHIEVQILGDRTGEVHLYERDCS 240

Qy  258 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 317
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 300

Qy  318 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 377
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360

Qy  378 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 437
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420

Qy  438 AEFTVSGVAINIGFLRALLREEDFTSKRIATCFIGDHPHLLQAPPADDEQGRILDYLADV 497
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADV 480

Qy  498 TVNKPHGVRPKDVAAPIDKLPLPRGSRDLKQLGPAAFARDLREQDALAVTDTT 557
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TVNKPHGVRPKDVAAPIDKLPLPRGSRDLKQLGPAAFARDLREQDALAVTDTT 540

Qy  558 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 617
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600

Qy  618 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 677
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 660

Qy  678 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR 737
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR 720

Qy  738 PAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQGADAVDGASAPLSGTTSQPS 797
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 PAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQGADAVDGASAPLSGTTSQPS 780

Qy  798 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVVRHEIPGGQL 857
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVVRHEIPGGQL 840

Qy  858 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 917
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900

Qy  918 AADPQKYDIPDSVIAFLRGELGNPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDA 977
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 AADPQKYDIPDSVIAFLRGELGNPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDA 960

Qy  978 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1037
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020

Qy 1038 LLVRLDAISEPDDKGMRNVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1097
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 LLVRLDAISEPDDKGMRNVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080

Qy 1098 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVPAATKVEGGDLIVVVS 1157
         ||||||||||||||||||||||||||||||:|||||||::|||||||||||||||||||
Db 1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVEGKIDRVVPAATKVEGGDLIVVVS 1140

RESULT 3
AAB01436
ID   AAB01436 standard; Protein; 1141 AA.
XX
AC   AAB01436;
XX
DT   20-OCT-2000  (first entry)
XX
DE   Pyruvate carboxylase of C. glutamicum.
```

```
XX
KW  Pyruvate carboxylase; expression; amino acid biosynthesis; lysine;
KW  glutamic acid; oxaloacetate; fermentation; biosynthesis.
XX
OS  Corynebacterium glutamicum.
XX
PN  WO20003905-A1.
XX
PD  06-JUL-2000.
XX
PF  23-DEC-1998;  98WO-US27301.
XX
PR  23-DEC-1998;  98WO-US27301.
XX
PA  (SINS/)  SINSKEY A J.
PA  (LESS/)  LESSARD P A.
PA  (WILL/)  WILLIS L B.
XX
PI  Sinskey AJ, Lessard PA, Willis LB;
XX
PT  WPI, 2000-465746/40.
DR  N-PSDB; AAA47533.
XX
PT  Novel polynucleotides encoding Corynebacterium glutamicum pyruvate
PT  carboxylase useful for industrial fermentation processes comprises a
PT  specific nucleotide sequence
XX
PS  Claim 3; Fig 1; 51pp; English.
XX
CC  The pyruvate carboxylase of Corynebacterium glutamicum can be used
CC  for producing amino acids, preferably lysine and glutamic acid in
CC  industrial fermentations and for replenishing oxaloacetate consumed
CC  for biosynthesis during growth. By incorporating the pyruvate
CC  carboxylase gene in expression vectors levels of expression can be
CC  2 - 20 fold higher than in Corynebacterium glutamicum.
XX
SQ  Sequence  1141 AA;

Query Match     98.3%;    Score 5764;    DB 21;    Length 1141;
Best Local Similarity   99.4%;    Pred. No. 0;
Matches 1134;   Conservative 5;   Mismatches 2;   Indels 0;   Gaps 0;

Qy    17  LVSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAV   76
          :|||||||||||:|||||||||||||||||||||||||:|||||||||||||||||||||
Db     1  MVSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAV   60

Qy    77  RIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPE  136
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  RIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPE  120

Qy   137  VLDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQIYPIFVKAVAGGGRGMR   196
          |||||||||||||||||||||||||||||||||:||||||||:|||||||||||||||||
Db   121  VLDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMR   180

Qy   197  FVSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDC  256
          ||:||||||||||||||||||||||:|||||||||||||||||||||:||||||||||||
Db   181  FVASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDC  240

Qy   257  SLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEM  316
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  SLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEM  300

Qy   317  NPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNN  376
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  NPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNN  360

Qy   377  GFRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRA  436
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  GFRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRA  420

Qy   437  LAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLAD  496
          |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db   421  LAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLAD  480

Qy   497  VTVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDT  556
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  VTVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDT  540

Qy   557  TFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLD  616
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  TFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLD  600

Qy   617  ELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRP  676
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  ELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRP  660

Qy   677  AIDAVLETNTAVAEVMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLL   736
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  AIDAVLETNTAVAEVMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLL   720

Qy   737  RPAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQP   796
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  RPAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQP   780

Qy   797  SLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQ   856
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  SLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQ   840

Qy   857  LSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPAD  916
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  LSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPAD  900

Qy   917  FAADPQKYDIPDSVIAFLRGELGNPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLD   976
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  FAADPQKYDIPDSVIAFLRGELGNPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLD   960

Qy   977  ADDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREEFYGLVEGRETLIRLPDVRT 1036
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  ADDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREEFYGLVEGRETLIRLPDVRT 1020

Qy  1037  PLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPF 1096
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  PLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPF 1080

Qy  1097  AGVTVTTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVVPAATKVEGGDLIVVV 1156
          |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db  1081  AGVTVTTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVV 1140

Qy  1157  S 1157
          |
Db  1141  S 1141

RESULT 4
AAG90511
ID  AAG90511 standard; Protein; 1140 AA.
XX
AC  AAG90511;
XX
DT  26-SEP-2001  (first entry)
XX
DE  C glutamicum protein fragment SEQ ID NO: 4265.
XX
KW  Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW  organic acid synthesis.
XX
OS  Corynebacterium glutamicum.
XX
PN  EP1108790-A2.
XX
PD  20-JUN-2001.
XX
PF  18-DEC-2000;  2000EP-0127688.
XX
PR  16-DEC-1999;  99JP-0377484.
PR  07-APR-2000;  2000JP-0159162.
```

```
PR   03-AUG-2000; 2000JP-0280988.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI   Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI   Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR   WPI; 2001-376931/40.
DR   N-PSDB; AAH55730.
XX
PT   Novel polynucleotides derived from Coryneform bacteria, for identifying
PT   mutation point of a gene, measuring expression of a gene, analysing
PT   expression profile or pattern of a gene and identifying homologous gene
PS   Claim 17; SEQ ID NO: 4265; 246pp + Sequence Listing; English.
XX
CC   The present invention provides a number of nucleotide and protein
CC   sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC   are useful for identifying the mutation point of a gene derived from a
CC   mutant of coryneform bacterium, measuring expression amount and
CC   analysing the expression profile or expression pattern of a gene derived
CC   from coryneform bacterium, and identifying a homologue of a gene derived
CC   from coryneform bacterium. Coryneform bacteria are useful for producing
CC   amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC   particularly L-lysine. The present sequence is a protein described
CC   in the exemplification of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   European Patent Office.
XX
SQ   Sequence   1140 AA;

Query Match       98.2%;   Score 5759;   DB 22;   Length 1140;
Best Local Similarity    99.4%;   Pred. No. 0;
Matches 1133;   Conservative    5;   Mismatches   2;   Indels   0;   Gaps   0;

Qy   18 VSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 77
         ::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60

Qy   78 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGVGFLSENAQLARECAENGITFIGPTPEV 137
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGVGFLSENAQLARECAENGITFIGPTPEV 120

Qy  138 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRF 197
         ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db  121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180

Qy  198 VSSPDELRKLATEASREAAAFGDGSVVERAVINPQHIEVQILGDRTGEVVHLYERDCS 257
         |:|||||||||||||||||||||||:||||||||||||||||||||:|||||||||||||
Db  181 VASPDELRKLATEASREAAAFGDGAVVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240

Qy  258 LQRRHQKVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 317
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 LQRRHQKVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 300

Qy  318 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 377
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360

Qy  378 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 437
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420

Qy  438 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADV 497
         ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db  421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV 480

Qy  498 TVNKPHGVRPKDVAAPIDKLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 557
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 540

Qy  558 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 617
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600

Qy  618 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 677
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 660

Qy  678 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR 737
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR 720

Qy  738 PAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 797
         ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 PAAVIKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 780

Qy  798 LSAIVAAFAHTPRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 857
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 LSAIVAAFAHTPRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840

Qy  858 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 917
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900

Qy  918 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA 977
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA 960

Qy  978 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLLRLPDVRTP 1037
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLLRLPDVRTP 1020

Qy 1038 LLVRLDAISEPDDKGMRNVVANVNGQIRPMVRDRSVESVTATAEKADSSNKGHVAAPFA 1097
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 LLVRLDAISEPDDKGMRNVVANVNGQIRPMVRDRSVESVTATAEKADSSNKGHVAAPFA 1080

Qy 1098 GVTIVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVPAATKVEGGDLIVVVS 1157
         |||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db 1081 GVTIVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVPAATKVEGGDLIVVVS 1140
```

```
PT   glutamic acid production in industrial fermentations -
XX
PS   Claim 1; Column 31-36; 29pp; English.
XX
CC   The present invention provides the protein and coding sequences of the
CC   Corynebacterium glutanicum pyruvate carboxylase protein. This is an
CC   enzyme in the anaplerotic pathway. It can be used in the replenishment of
CC   oxaloacetate consumed during lysine and glutamic acid production in
CC   industrial fermentation.
XX
SQ   Sequence   1140 AA;

Query Match          98.2%;   Score 5759;   DB 22;   Length 1140;
Best Local Similarity  99.4%;   Pred. NO. 0;
Matches 1133;   Conservative  5;   Mismatches  2;   Indels  0;   Gaps  0;

Qy   18  VSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR   77
         :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR   60

Qy   78  IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  137
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  120

Qy  138  LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRF  197
         |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db  121  LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTPIFVKAVAGGGGRGMRF   180

Qy  198  VSSPDELRKLATEASREAAAFGDGSVYVERAVINPQHEVQILGDRTGEVVHLYERDCS   257
         :|||||||||||||||||||||:|||||||||||||||||||:||||||||||||||||
Db  181  VASPDELRKLATEASREAAAFGDGAVYVERAVINPQHEVQILGDHTGEVVHLYERDCS   240

Qy  258  LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN  317
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN  300

Qy  318  PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG  377
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG  360

Qy  378  FRPDTGITTAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  437
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  FRPDTGITTAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  420

Qy  438  AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADV  497
         ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db  421  AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV  480

Qy  498  TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVIDTT  557
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db  481  TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  540

Qy  558  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  617
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  600

Qy  618  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIPDALNDVSQMRPA  677
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIPDALNDVSQMRPA  660

Qy  678  IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR  737
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR  720

Qy  738  PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  797
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  780

Qy  798  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL  857
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL  840

Qy  858  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF  917
```

```
Db  841  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF  900

Qy  918  AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA  977
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA  960

Qy  978  DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALLDREFYGLVEGRETLIRLPDVRTP  1037
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALLDREFYGLVEGRETLIRLPDVRTP  1020

Qy 1038  LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA  1097
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA  1080

Qy 1098  GVVTVTVAEGDEVKAGDAVAIIEAMKMEATTTASVDGKIERVVPAATKVEGGDLIVVVS  1157
         |||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db 1081  GVVTVTVAEGDEVKAGDAVAIIEAMKMEATTTASVDGKIDRVVVPAATKVEGGDLIVVVS  1140

RESULT 6
AAE25601
ID   AAE25601 standard; Protein; 1140 AA.
XX
AC   AAE25601;
XX
DT   04-NOV-2002  (first entry)
XX
DE   Corynebacterium glutamicum pyruvate carboxylase protein.
XX
KW   Pyruvate carboxylase; anaplerotic enzyme; industrial fermentation;
KW   oxaloacetate; growth; enzyme.
XX
OS   Corynebacterium glutamicum.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "Encoded by GTG"
XX
PN   US6403351-B1.
XX
PD   11-JUN-2002.
XX
PF   03-OCT-2000;  2000US-0677575.
XX
PR   23-DEC-1998;  98US-0220081.
XX
PA   (ARCH ) ARCHER-DANIELS MIDLAND CO.
XX
PI   Sinskey AJ,  Lessard PA,  Willis LB;
XX
DR   WPI; 2002-536037/57.
DR   N-PSDB; AAD42046.
XX
PT   Novel pyruvate carboxylase polypeptide, useful for replenishing
PT   oxaloacetate consumed for biosynthesis during growth, or lysine and
PT   glutamic acid production in industrial fermentation -
XX
PS   Claim 1; Column 29-36; 28pp; English.
XX
CC   The present invention relates to novel pyruvate carboxylase proteins and
CC   polynucleotides encoding such proteins. Sequences of the invention are
CC   important anaplerotic enzymes for replenishing oxaloacetate consumed for
CC   biosynthesis during growth, or lysine and glutamic acid production in
CC   industrial fermentation. The present sequence is C. glutamicum pyruvate
CC   carboxylase protein.
XX
SQ   Sequence   1140 AA;

Query Match          98.2%;   Score 5759;   DB 23;   Length 1140;
Best Local Similarity  99.4%;   Pred. No. 0;
Matches 1133;   Conservative  5;   Mismatches  2;   Indels  0;   Gaps  0;
```

```
Db    1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS 1140
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
RESULT 7
AAU98053
ID    AAU98053 standard; Protein; 1140 AA.
XX
AC    AAU98053;
XX
DT    27-AUG-2002 (first entry)
XX
DE    Corynebacterium wild-type feedback-resistant pyruvate carboxylase enzyme.
XX
KW    Feedback-resistant; pyruvate carboxylase; enzyme;
KW    aspartic acid feedback inhibition resistant.
XX
OS    Corynebacterium glutamicum.
XX
XX
PN    WO200231158-A2.
XX
PD    18-APR-2002.
XX
PF    12-OCT-2001; 2001WO-US31893.
XX
PR    13-OCT-2000; 2000US-239913P.
XX
PA    (ARCH ) ARCHER-DANIELS MIDLAND CO.
XX
PI    Hanke PD;
XX
DR    WPI; 2002-463267/49.
XX
PT    Novel mutated, feedback resistant pyruvate carboxylase enzyme
PT    polypeptide, useful for producing amino acids e.g. L-lysine,
PT    L-threonine, L-glycine, L-glutamic acid, L-proline and L-methionine and
PT    L-isoleucine. -
XX
PS    Disclosure; Fig 2; 42pp; English.
XX
CC    The present invention relates to a new mutated, feedback-resistant
CC    pyruvate carboxylase enzyme. The invention is useful for producing an
CC    amino acid (e.g. L-Lys, L-Thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro),
CC    by culturing a host cell in a suitable media and separating the amino
CC    acid from the medium. The vector of the invention is useful for
CC    replacement of a wild-type pyruvate carboxylase gene, with a feedback
CC    resistant pyruvate carboxylase gene, in a Corynebacterium glutamicum, by
CC    replacing a genomic copy of the wild-type pyruvate carboxylase gene with
CC    a selectable marker gene through homologous recombination to form a first
CC    recombination strain, and replacing the selectable marker gene in the
CC    first recombinant strain, with feedback resistant pyruvate carboxylase
CC    gene through homologous recombination to form a second recombinant
CC    strain, where the homologous recombination in the above steps, occurs
CC    between the host cell and the vector. The feedback-resistant pyruvate
CC    carboxylase enzyme is resistant to feedback inhibition from aspartic
CC    acid. The present amino acid sequence represents the wild-type
CC    feedback-resistant pyruvate carboxylase enzyme of the invention.
XX
SQ    Sequence    1140 AA;

Query Match         98.2%;  Score 5759;  DB 23;  Length 1140;
Best Local Similarity   99.4%;
Matches 1133;  Conservative   5;  Mismatches   2;  Indels   0;  Gaps   0;

Qy    18 VSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 77
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60

Qy    78 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 137
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120

Qy   138 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 197
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180
```

```
XX
DT   30-JUN-1999 (first entry)
XX
DE   C. glutamicum pyruvate carboxylase protein.
XX
KW   Pyruvate carboxylase; amino acid production; lysine production;
KW   threonine production; homoserine production; glutamate production;
KW   arginine production; feed additive; condiment; pharmaceutical;
KW   fine chemical; ss.
XX
OS   Corynebacterium glutamicum.
XX
PN   DE19831609-A1.
XX
PD   15-APR-1999.
XX
PF   14-JUL-1998;   98DE-1031609.
XX
PR   04-OCT-1997;   97DE-1043894.
XX
PA   (KERJ ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX
PI   Eikmanns B,  Peters-Wendisch P,  Sahm H;
XX
DR   WPI; 1999-245521/21.
DR   N-PSDB; AAX24102.
XX
PT   Increasing microbial production of specific amino acids by
PT   increasing activity or expression of pyruvate carboxylase
XX
PS   Claim 15; Page 11-15; 18pp; German.
XX
CC   This invention describes the isolation of a pyruvate carboxylase
CC   from Corynebacterium glutamicum which is used in a novel method for
CC   production of lysine, threonine, homoserine, glutamate and/or arginine,
CC   variously useful as feed additives, condiments, pharmaceuticals and
CC   intermediates for fine chemicals. Increasing pyruvate carboxylase
CC   activity increases the yield of microbial production of amino acids
CC   of the asparate and/or glutamate families, e.g. about 50% more lysine,
CC   40% more threonine and 150% more homoserine are secreted into the
CC   culture medium.
XX
SQ   Sequence   1140 AA;

Query Match     98.1%;   Score 5753;   DB 20;   Length 1140;
Best Local Similarity   99.3%;   Pred. No. 0;
Matches 1132;   Conservative   5;   Mismatches   3;   Indels   0;   Gaps   0;

Qy    18  VSTHTSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR   77
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db     1  MSTHTSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR   60

Qy    78  IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  137
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  120

Qy   138  LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRF  197
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRF  180

Qy   198  VSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCS  257
          |:||||||||||||||||||||||||:|||||||||||||||| |||||||||||||||
Db   181  VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS  240

Qy   258  LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN  317
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN  300

Qy   318  PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG  377
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG  360

Qy   378  FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  437
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  420

Qy   438  AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADV  497
          ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db   421  AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV  480

Qy   498  TVNKPHGVRPKDVAAPIDKLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  557
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  TVNKPHGVRPKDVAAPIDKLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  540

Qy   558  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGATYDVAMRFLFEDPWDRLDE  617
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGATYDVAMRFLFEDPWDRLDE  600

Qy   618  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  677
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  660

Qy   678  IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTIDYYIKMAEEIVKSGAHILAIKDMAGLLR  737
          ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db   661  IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYIKMAEEIVKSGAHILAIKDMAGLLR  720

Qy   738  PAAVTKLVTALRREFDLPVHVTHTDTAGGQLATYPAAAQAGADAVDGASAPLSGTTSQPS  797
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  PAAVTKLVTALRREFDLPVHVTHTDTAGGQLATYPAAAQAGADAVDGASAPLSGTTSQPS  780

Qy   798  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFFESGTPGPTGRVYRHEIPGGQL  857
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFFESGTPGPTGRVYRHEIPGGQL  840

Qy   858  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF  917
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF  900

Qy   918  AADPQKYDIPDSVIAFLRGELGNPPGGWEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA  977
          ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db   901  AADPQKYDIPDSVIAFLRGELGNPPGGWEPLRTRALEGRSEGKAPLITEVPEEEQAHLDA  960

Qy   978  DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP  1037
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP  1020

Qy  1038  LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA  1097
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA  1080

Qy  1098  GVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVPAATKVEGGDLIVVVS  1157
          ||||||||||||||||||||||||||||||||||||||:||||:||||||||||||||
Db  1081  GVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS  1140
```

RESULT 8
AAW93971
ID   AAW93971 standard; Protein; 1140 AA.
XX
AC   AAW93971;

```
Query Match             98.1%;   Score 5751;   DB 22;   Length 1140;
Best Local Similarity   99.3%;   Pred. No. 0;
Matches 1132;   Conservative   5;   Mismatches   3;   Indels   0;   Gaps   0;

Qy     18 VSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR   77
           :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR   60

Qy     78 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  137
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  120

Qy    138 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRF  197
           ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db    121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNDEIVKSAEGQTYPIFVKAVAGGGRGMRF  180

Qy    198 VSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCS  257
           |:|||||||||||||||||||||:|||||||||||||||||||:|||||||||||||||
Db    181 VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS  240

Qy    258 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN  317
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN  300

Qy    318 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATIKELGLTQDKIKTHGAALQCRITTEDPNNG  377
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATIKELGLTQDKIKTHGAALQCRITTEDPNNG  360

Qy    378 FRPDTGTITAYRSPGGAGVRLDGAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  437
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 FRPDTGTITAYRSPGGAGVRLDGAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  420

Qy    438 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADV  497
           ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db    421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV  480

Qy    498 TVNKPHGVRPKDVAAPIDKLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  557
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TVNKPHGVRPKDVAAPIDKLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  540

Qy    558 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  617
           |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db    541 FRDAHQSLLATRVRSFALKPAAEAVAKKTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  600

Qy    618 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  677
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  660

Qy    678 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR  737
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR  720

Qy    738 PAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  797
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 PAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  780

Qy    798 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQL  857
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQL  840

Qy    858 SNLRAQAFALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF  917
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 SNLRAQAFALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF  900

Qy    918 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA  977
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA  960

Qy    978 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1037
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020

Qy   1038 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1097
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080

Qy   1098 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVVPAATKVEGGDLIVVVS 1157
           |||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db   1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS 1140
```

```
RESULT 9
AAG93249
ID   AAG93249 standard; Protein; 1140 AA.
XX
AC   AAG93249;
XX
DT   26-SEP-2001   (first entry)
XX
DE   C glutamicum protein fragment mutant P458S.
XX
KW   Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW   organic acid synthesis; mutant; mutein.
XX
OS   Corynebacterium glutamicum.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 458
FT                   /note= "wild-type Pro substituted by Ser"
XX
PN   EP1108790-A2.
```

```
PA   (AJIN ) AJINOMOTO CO INC.
XX
PI   Hirano S, Nonaka G, Matsuzaki Y, Akiyoshi N, Nakamura K, Kimura E;
PI   Osumi T, Matsui K, Kawahara Y, Kurahashi O, Nakamatsu T;
PI   Sugimoto S;
XX
DR   WPI; 2001-300170/31.
DR   N-PSDB; AAF87437.
XX
PT   Proteins and their DNA useful for microbial production of L-amino acids
PT   -
XX
PS   Claim 9; Page 132-135; 215pp; Japanese.
XX
CC   The present sequence is provided in a specification relating to genes
CC   encoding thermophilic amino acid biosynthesis system enzymes of
CC   the thermotolerant bacterium Corynebacterium thermoaminogenes.
CC   The novel proteins retain at least 30% isocitrate ligase activity
CC   after heating at 50oC for 5 minutes. DNA fragments encoding the
CC   enzymes were isolated from a Corynebacterium thermoaminogenes
CC   chromosomal DNA plasmid library by PCR. The DNA may be used for
CC   developing strains of amino acid producing microorganisms.
XX
SQ   Sequence  1139 AA;

Query Match             90.4%;  Score 5299.5;  DB 22;  Length 1139;
Best Local Similarity   90.9%;  Pred. No. 0;
Matches 1036;  Conservative  44;  Mismatches  59;  Indels  1;  Gaps  1;

Qy     18 VSTHTSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 77
           | ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 VVTTTPSTLPAFKKILVANRGEIAVRAFRAAYETGAATVAIYPREDRGSFHRSFASEAVR 60

Qy     78 IGTEGSPVKAYLDIDEIIGAAKKVKADAITPGYGFLSENAQLARECAENGITFIGPTPEV 137
           |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db     61 IGTEGSPVKAYLDIDEIINAAKKVKADAVYPGYGFLSENAQLARECAENGITFIGPTPEV 120

Qy    138 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRF 197
           |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||:
Db    121 LDLTGDKSKAVSAAKKAGLPVLAESTPSTDIEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180

Qy    198 VSSPDELRKLATEASREAAAFGDGSVYERAVINPQHIEVQILGDRTGEVVHLYERDCS 257
           |::|::||||||||||||||||||||||||||||||||||||||:||:||||||||||
Db    181 VEKPEDLRELAREASREAAAFGDGSVYERAVIKPQHIEVQILGDHTGDVIHLYERDCS 240

Qy    258 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN 317
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCKSIGYQAGTVEFLVDEAGNHVFIEMN 300

Qy    318 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGTQDKIKTHGAALQCRITTEDPNNG 377
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 PRIQVEHTVTEEVTSVDLVKAQMHLAAGAILKELGLTQDKITTHGAALQCRITTEDPSNN 360

Qy    378 FRPDIGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 437
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 FRPDTGVITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVSRAQRAL 420

Qy    438 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILLYDLADV 497
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AEFNVSGVATNIGFLRALLREEDFTKRRIDTGFIGSHQHLLQAPPADDEQGRILEYLADV 480

Qy    498 TVNKPHGCVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 557
           |||||||| :::||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TVNKPHGERP-ETARPIEKLPEVENIPLPRGSRDRLKQLGPEGFARDLREQDALAVTDTT 539

Qy    558 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 617
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 FRDAHQSLLATRVRSFALTPAARAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWARLDE 599

Qy    618 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 677
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVQEAAKSGVDIFRIFDALNDISQMRPA 659
```

```
Db    421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHSHLLQAPPADDEQGRILDYLADV 480

Qy    498 TVNKPHGCVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 557
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TVNKPHGCVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 540

Qy    558 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 617
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600

Qy    618 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 677
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 660

Qy    678 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYILKMAEEIVKSGAHILAIKDMAGLLR 737
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYILKMAEEIVKSGAHILAIKDMAGLLR 720

Qy    738 PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 797
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 780

Qy    798 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQL 857
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQL 840

Qy    858 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 917
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900

Qy    918 AADPQKYDIPDSVIAFLRGELGNPPGGWEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA 977
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AADPQKYDIPDSVIAFLRGELGNPPGGWEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA 960

Qy    978 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1037
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020

Qy   1038 LLVRLDAISEPDKGMRNVVANVNGQIRPMRVRDRSVESVTAEKADSSNKGHVAAPPA 1097
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 LLVRLDAISEPDKGMRNVVANVNGQIRPMRVRDRSVESVTATEKADSSNKGHVAAPPA 1080

Qy   1098 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVPAAIKVEGGDLIVVVS 1157
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAAIKVEGGDLIVVVS 1140
```

```
RESULT 10
AAB83180
ID   AAB83180 standard; Protein; 1139 AA.
XX
AC   AAB83180;
XX
DT   09-JUL-2001 (first entry)
XX
DE   Corynebacterium thermoaminogenes pc protein.
XX
KW   Corynebacterium; thermophilic; amino acid biosynthesis; enzyme;
KW   thermotolerant; aceA; accBC; dtsR1; dtsR2; pfk; scrB; gluABCD;
KW   pdhA; pc; ppc; acn; icd; lpd; odhA.
XX
OS   Corynebacterium thermoaminogenes.
XX
PN   WO200125447-A1.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-JP06913.
XX
PR   04-OCT-1999; 99JP-0282716.
PR   01-NOV-1999; 99JP-0311147.
PR   21-APR-2000; 2000JP-0120687.
XX
```

```
Qy   678  IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR  737
          :|::|||||||  :|||::|| ||||||| :||:|| |||||||||||||||||||
Db   660  IDAVLETGTSVAEVAMAYSGDLSNPGEKLYTLDYYLNLAEQIYDSGAHILAIKDMAGLLR  719

Qy   738  PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  797
          ||| ||||||||||||||||||||||||||||| :||| |||||| ||||||||||||||
Db   720  RAAAPKLVTALRREFDLPVHVHTHDTAGGQLATYLAAANAGADAVDAASAPLSGTTSQPS  779

Qy   798  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL  857
          :||:||||||||||||:|:||||||||||||||||||||||||||:||||||||||||||
Db   780  MSALVAAFAHTRRDTGLNLQAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL  839

Qy   858  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF  917
          |||||| | |||||||||| |||||||||||||||||||||||||||||||||||| ||
Db   840  SNLRAQAVALGLADRFELIEDYYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVSPEDF  899

Qy   918  AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA  977
          ||||||||||||||||||||| |:|||||||||||||||||:|||| ||||||||||||
Db   900  AADPQKYDIPDSVIAFLRGELGTPPGGWPEPLRTRALEGRSQGKAPLAEIPAEEQAHLDS  959

Qy   978  DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP  1037
          |||| || ||||||||||||||||||||||:|||||||||||||||| || |:|||||||
Db   960  DDSAEERGTLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLKEGREELIRLTGVSTP  1019

Qy  1038  LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA  1097
          ::|||||| ::::|||||||:||||||||||||||||:|:||||||||:::||||||||
Db  1020  MVVRLDAVSEPDDKGMRNVVNVNGQIRPKVRDRSVESVTATAEKADATNKGHVAAPFA   1079

Qy  1098  GVVTVTVAEGDEVKAGDAVAIIEAMKMEATTTASVDGKIERVVVPAATKVEGGDLIVVVS  1157
          |||||||||||:||||:|||||||||||||||||||||||:|||||||||||||||||||
Db  1080  GVVTVTVAEGDEIKAGDAVAIIEAMKMEATTTAPVDGKIERVVVPAATKVEGGDLIVVVS  1139
```

RESULT 11
AAB79302
```
ID   AAB79302 standard; Protein; 532 AA.
XX
AC   AAB79302;
XX
DT   30-APR-2001  (first entry)
XX
DE   Corynebacterium glutamicum SMP protein sequence SEQ ID NO:120.
XX
KW   Corynebacterium glutamicum; carbon metabolism and energy production;
KW   SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW   fine chemical production; organic acid; proteinogenic amino acid;
KW   nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW   nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW   carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW   diagnosis; Corynebacterium diphtheriae; evolutionary study.
XX
OS   Corynebacterium glutamicum.
XX
PN   WO200100844-A2.
XX
PD   04-JAN-2001.
XX
PF   23-JUN-2000; 2000WO-IB00943.
XX
PR   25-JUN-1999;   99US-0141031.
PR   08-JUL-1999;   99DE-1031412.
PR   08-JUL-1999;   99DE-1031413.
PR   08-JUL-1999;   99DE-1031419.
PR   08-JUL-1999;   99DE-1031420.
PR   08-JUL-1999;   99DE-1031424.
PR   08-JUL-1999;   99DE-1031428.
PR   08-JUL-1999;   99DE-1031431.
PR   08-JUL-1999;   99DE-1031433.
PR   08-JUL-1999;   99DE-1031434.
PR   08-JUL-1999;   99DE-1031510.
PR   08-JUL-1999;   99DE-1031562.
PR   08-JUL-1999;   99DE-1031634.
PR   09-JUL-1999;   99DE-1032180.
PR   09-JUL-1999;   99DE-1032227.
PR   09-JUL-1999;   99DE-1032230.
PR   09-JUL-1999;   99US-0143208.
PR   14-JUL-1999;   99DE-1032924.
PR   14-JUL-1999;   99DE-1032973.
PR   14-JUL-1999;   99DE-1033005.
PR   27-AUG-1999;   99DE-1040765.
PR   31-AUG-1999;   99US-0151572.
PR   03-SEP-1999;   99DE-1042076.
PR   03-SEP-1999;   99DE-1042079.
PR   03-SEP-1999;   99DE-1042086.
PR   03-SEP-1999;   99DE-1042087.
PR   03-SEP-1999;   99DE-1042088.
PR   03-SEP-1999;   99DE-1042095.
PR   03-SEP-1999;   99DE-1042123.
PR   03-SEP-1999;   99DE-1042125.
XX
PA   (BADI ) BASF AG.
XX
PI   Pompejus M,  Kroeger B,  Schroeder H,  Zelder O,  Haberhauer G;
XX
DR   WPI; 2001-061975/07.
DR   N-PSDB; AAF71419.
XX
PT   New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
PT   metabolism and oxidative phosphorylation protein for production or
PT   modulation of production of fine chemicals e.g. amino acids,
PT   carbohydrates or enzymes -
XX
PS   Claim 20; Page 319-320; 1246pp; English.
XX
CC   AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
CC   metabolism and oxidative phosphorylation (SMP) proteins given in
CC   AAB79243 to AAB 79633 which are involved in carbon metabolism and
CC   energy production. The C. glutamicum SMP gene can be used in vectors
CC   (II) for expression in host cells and production or modulation of
CC   production of fine chemicals, such as, an organic acid, a proteinogenic
CC   or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,
CC   a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty
CC   acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
CC   cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins
CC   (III) encoded by them are used for diagnosing the presence or activity of
CC   Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
CC   containing them are used to map genomes of organisms related to
CC   C. glutamicum, identify and localise C. glutamicum sequences of interest,
CC   in evolutionary studies, in determining SMP protein regions required
CC   for function, in modulating SMP protein activity, in modulating the
CC   metabolism of sugars, and in modulating high-energy molecule production
CC   in a cell (i.e. ATP, NADPH).
XX
SQ   Sequence  532 AA;

Query Match             45.7%;  Score 2678;  DB 22;  Length 532;
Best Local Similarity   98.9%;  Pred. No.7.5e-191;
Matches  526;   Conservative    4;  Mismatches   2;   Indels   0;   Gaps   0;
```

```
Qy     1  MTAITLGGLLKGIITILVSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYP  60
          :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  VTAITLGGLLKGIITILVSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYP  60

Qy    61  REDRGSFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLA  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  REDRGSFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLA  120

Qy   121  RECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDIVKSAEGQTY  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  RECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTY  180

Qy   181  PIFVKAVAGGGGRGMRFVSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQI  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  PIFVKAVAGGGGRGMRFVASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQI  240

Qy   241  LGDRTGEVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGT  300
```

```
Best Local Similarity 46.0%;  Pred. No. 2.5e-179;
Matches  533;  Conservative  188;  Mismatches  390;  Indels  48;  Gaps  13;

Qy   28  AFKKILVANRGEIAVRAFRAALETGAATVAIIYPREDRGSFHRSFASEAVRICTEGSPVKA  87
         :::|||||||||||||| ||:|| ||  :|  :|  |: : :|| |:| |  |:  | |::|
Db    5  SIQKVLVANRGEIAIRIFRACTELNIRTVAVYSKEDSGSYHRYKADEAYLVGEGKKPIDA   64

Qy   88  YLDIDEIIGAKKVKADAIPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRA    147
         ||||  || |:| ||||:||||||||| : || |  :||.|||| :|  ||| ||| |
Db   65  YLDIEGIIDIAKRNKVDAIHPGYGFLSENIHFARRCEEEGIVFIGPKSEHLDMFGDKVKA  124

Qy  148  VTAAKAGLPVLAEST-PSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRK   206
         . |||||:||: . :    || .   |: | | ||:||| ||||||||||||::::| :|
Db  125  REQAEKAGIPVIPGSDGPAETLEAVEQFGQANGYPIIIKASLGGGGRGMRIVRSESEVKE  184

Qy  207  LATEASREAAFGDGSVVYERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVV    266
         . | |: |:|| :. |:|: ::| | ||||| ||::.  |||:||||| ||||||||:
Db  185  AYERAKSEAKAFGNDEVYVEKLIENPKHIEVQVIGDKQGNVHLFERDCSVQRRHQKVI   244

Qy  267  EIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTV  326
         |:||:   |.|||| ||  ||  |:::.| :|||||||| .|||: ||| ||||:|||||
Db  245  EVAPSVSLSPELRDQICEAAVALAKNVNYINAGTVEFLV-ANNEFYFIEVNPRVQVEHTI  303

Qy  327  TEEVTEVDLVKAQMRLAAGATL--KELGLTQDK-IKTHGAALQCRITTEDPNNGFRPDTG  383
         ||.:| |:|:   :: ||  :|  |::.: : | |.| | | |  ||||||:|.| |||
Db  304  TEMITGVDIVQTQILVAQGHSLHSKKVNIPEQKDIFTIGYAIQSRVTTEDPQNDFMPDTG  363

Qy  384  TITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTV  442
         .| |||| |. |||| |.: ||| .|| :.||||::| . | :||: | ::  |:| |:|
Db  364  KIMAYRSGGGFGVRLDTGNSFQGAVITPYDSLLVKLSTWALTFEQAAAKMVRNLQEFRI   423

Qy  443  SGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVN--  500
         .|: ||| || |. ||::| . ::   || : | | |  |:|  || |  :|  |||
Db  424  RGIKTNIPFLENVAKHEKFLTGQYDTSFIDTTPELFNFPKQKDRGTKMLTYIGNVTVNGF  483

Qy  501  ------KPHGVRPKDVAAPIDKLPLPRGSRDRLKQLGPAAFARDLREQDALAV         553
               |.   .|:|.   :|| .  .  :  ::  :|. ||  :  :||: |::
Db  484  PGIGKKEKPAFDKPLGVKVDVDQQP-------ARGTKQILDEKGAEGLANWVKEQKSVLL  536

Qy  554  TDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWD  613
         |||||||||||||||||||:|: | ||  : | |:|| :| |||||:||| |||:|||||
Db  537  TDTTFRDAHQSLLATRIRSHDLKKIANPTAALWPELFSMEMWGGATFDVAYRFLKEDPWK  596

Qy  614  RLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKPAASSGVDIPRIFDALNDVSQ  673
         || :|| : ||.. |:|||. |:|||| |||.| : |||.|: |||:: |||| ||::  |
Db  597  RLEDLRKEVPNTLFQMLLRSSNAVGYTNYPDNVIKEFVKQSAQSGIDVFRIFDSLNWVKG  656

Qy  674  MRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMA  733
         | .||||| :|. .|||:|   :|| .| |  || |:||| ||||:| |||||:|||||
Db  657  MTLAIDAVRDTG-KVAEAAICYTGDILDKNRTKYDLAYYTSMAKELEAAGAHILGIKDMA  715

Qy  734  GLLRPAAVTKLVTALRREFDLPVHVTHTDTAGGQLATYFAAAQAGADAVDGASAPLSGTT  793
         ||| | |  :|| || ||.::||| .|||| ||:  .: | | || :|::. | :. ::|
Db  716  GLLKPQAAYELVSALKETIDIPVHLHTHDTSGNGIYMAKAVEAGVDIIDVAVSSMAGLT   775

Qy  794  SQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIP   853
         |||| |:  |:    :::.  : : | ::  |||:|| |: |::||   || | :.|.:|
Db  776  SQPSASGFYHAMEGNDRRPEMNVQGVELLSQYWESVRKYSEFESGMKSPHTEYEHEMP   835

Qy  854  GGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVD  913
         |||:|||   |  :|||||: |:|:|  .|||::   |:|||||||||||:||::|:.. |
Db  836  GGQYSNLQQQAKGVGLGDRKNEVKEMYRRVNDMFGDIVKTPSSKVVGDMALYMVQNNLT   895

Qy  914  PADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEE---  970
               . : || |||||:  ||:|:|:|  ||:|| | ::  |    .| :  :.
Db  896  EKDVYEKGESLDFPDSVVELFKGNIGQPHGGFPEKLQKLILKGQE---PITVRPGELLE   951

Qy  971  ------------EQAHLDADDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFF 1018
                     : || |:  |   :|||| .|:  | :  : | |:|:|:|:::
Db  952  PVSFEAIKQEFKEQHLEISD----QDAVAYALYPKVFDYVKTESYGDISVLDTPFF    1007

Qy 1019  YGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIREPMVRDRSVESVT 1078
```

RESULT 12

AAU00511

```
ID   AAU00511    standard; Protein;  1148 AA.
XX
AC   AAU00511;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Bacillus subtilis pyruvate carboxylase enzyme A.
XX
KW   Pyruvate carboxylase A; PycA; oxaloacetate; Escherichia;
KW   L-amino acid production; fermentation.
XX
OS   Bacillus subtilis strain 168.
XX
PN   EP1092776-A1.
XX
PD   18-APR-2001.
XX
PF   05-OCT-2000; 2000EP-0121763.
XX
PR   14-OCT-1999;  99RU-0121636.
XX
PA   (AJIN ) AJINOMOTO CO INC.
XX
PI   Gusyatiner MM, Kozlov YI, Ptitsyn LR, Altman IB, Voroshilova EB;
PI   Iomantas YAV, Yampolskaya TA;
XX
DR   WPI; 2001-309819/33.
DR   N-PSDB; AAS01509.
XX
PT   New bacterium from the genus Escherichia containing a gene encoding for
PT   pyruvate carboxylase useful for producing higher concentrations of
PT   L-amino acids -
XX
PS   Claim 4; Page 17-21; 28pp; English.
XX
CC   The present sequence represents Bacillus subtilis pyruvate
CC   carboxylase enzyme A. The PycA gene encodes for pyruvate carboxylase
CC   which catalyses the carboxylation of pyruvate to form oxaloacetate.
CC   Transformation of the Bacillus subtilis pycA gene into a bacterium
CC   belonging to the genus Escherichia results in the bacterium showing
CC   L-amino acid productivity. The invention provides a novel method for
CC   producing an L-amino acid by fermentation. The method involves
CC   cultivating the bacterium in a medium and producing and accumulating
CC   the L-amino acid in the medium. The new bacterium harbouring the gene
CC   coding for an enzyme having pyruvate carboxylase activity is useful
CC   for producing higher concentrations of L-amino acids in vitro than
CC   prior art.
XX
SQ   Sequence  1148 AA;

Query Match    43.18%;  Score 2530;  DB 22;  Length 1148;
```

```
Db  241  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         LGDHTGEVVHLYERDCSLQRRHQKVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGT   300

Qy  301  VEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTIQDKIKT 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  VEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTIQDKIKT 360

Qy  361  HGAALQCRITTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMT  420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  HGAALQCRITTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMT  420

Qy  421  CRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQA  480
         |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db  421  CRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQA  480

Qy  481  PPADDEQGRILDYLADVTVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDR         532
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  PPADDEQGRILDYLADVTVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDR         532
```

```
Query Match                42.5%;  Score 2494;  DB 23;    Length 1146;
Best Local Similarity      45.7%;  Pred. No. 1.2e-176;
Matches    525;  Conservative 186;  Mismatches 408;  Indels  30;  Gaps  11;

Qy     30 KKIVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASFAVRIGTEGSPVKAYL  89
          |||:|||||:|| |:|:||| ||  :||| ||||||:| |:| |: ::|||||||||
Db      5 KKVLVANRGEIAIRVMRACTELKIKTVAIYSQEDTGSFHRYKSDEAYLVGAGKKPIDAYL  64

Qy     90 DIDEIIGAAKKVKADAIYPGYGFLSENAQLAREECAENGITFIGPTPEVLLTGDKSRAVI  149
          |:| :| |||| |: ||||||||||||: | |  :||| :|| | :|||  |:|| ||:|
Db     65 DIENIEIAKESGADAIHPGYGFLSENIEFARRCEQEGIIFVGPKSKHLDMFGDKIKAKE  124

Qy    150 AAKKAGLPVLAEST-PSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKLA  208
          ||||||| |:| ||  ||:|||: :   ||||||::|||||||||||||| |:|:||||
Db    125 QALLADIPVIPGSNGPVAGIKEVEEFGEKNGYPLMIKASLGGGGRGMRVVESKEHVKESF  184

Qy    209 TEASREAEAAFGDGSVYERAVINPQHIEVQILGDRTGEVHLYERDCSLQRRHQKVVEI  268
          || : ||| ||| |: |||: ||||:||||||||  ||:| || :|:||||||||||
Db    185 ERASSEAKAAFGNDEVYVEKCVMNPKHIEVQILGDTHGNIVHLFERDCSIQRRHQKVVEV  244

Qy    269 APAQHLDPELRDRICADAVKFCRSIGYQQGAGTVEFLVDEKGNHYFIEMNPRIQVEHTVE  328
          || :||||| ||||   :|: |||||||:|  | ||| :| |||||||||||:| || :
Db    245 APCNAITSELRNRICDAAVKLMKNVDYINAGTVEFLV-EGDDFYFIEVNPRVQVEHTIFE  303

Qy    329 EVTEVDLVKAQMRLAAGATLKELGLT---QDKIKTHGAALQCRITTEDPNNGFRPDTGTI  385
          | :|:| | : :| :  ||::||:|:   |:|||||||:| |||||||||||:|||||||
Db    304 MITGIDIVQSQLFIADGYALHDQLVAIPKQEDIHHGSAIQSRITTEDPLNNFMPDTGRV  363

Qy    386 TAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSG  444
          | ||:  | | |||  ||| :|||  |||:|||: :||||| | : | | ||:|||:|
Db    364 DTYRSTGGFGVRLDAGNGFQGTVVTPFYDSLLVKLCTWGMTFEQATRKMRRNLIEFRIRG  423

Qy    445 VATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKPHG  504
          | ||| |:| |||||| :  | | ||||||| |  |:||||||| || ||| ||||| ||
Db    424 VKTNIPFLLNVVRHPDFASGNYNTSFIDTTPELFKFPHIRDRGTKTLRYIGNVTVNGFPG  483

Qy    505 VRPKDVAAPIDKLPNIKDLP----LPRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRD  560
          |||:|:||||||| |||| :    | |||: :| | ||| |  :||: |  |||||| ||
Db    484 IKHRD--KPVYAEPRLPKIPYGSQISPGTKQILDAKGPEGVVDWVKKQEEVLLTDTTLRD  541

Qy    561 AHQSLLATRVRSPALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELRE  620
          ||||||||||||| ||||||||| | || ||||||||||||| | |||:|||: |||:||
Db    542 AHQSLLATRVRSKDIFQVADAMAHLLPNMFSFEMWGGATFDVAYRFLNEDPWVRLETLRK  601

Qy    621 AMPNVNIQMLLRGRNTVGYTPYPDSVCRAPVKEAASSGVDIFRIFDALNDVSQMRPAIDA  680
          |:|||  ||||| |||||| ||||||  ||||| ||||||:||||||| ||| ||  ||
Db    602 QIPNVMFQMLLRGANAVGYKNYPDNVIREFVKQSAQSGVDVFRVFDSLNWIKGMEVSIDA  661

Qy    681 VLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAA  740
          |:| ||||:|| |  ||:|  | | |||||| ||||:||||| ||||| ||||||| ||
Db    662 VREAG-KVVEAAICYTGDIDDDTRTKYTIDYYKDMAKELVAQGTHILGIKDMAGLLKPQA  720

Qy    741 VTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSA  800
          | ||::||||||  |||:||||:| ||||||  ||:|||||| | ||:|::|  ||::
Db    721 AYRLIGELKDTVDVPIHLHTHDTSGNGLTYAAAVSAGVDIVDVASSAMSGATSQPSMTG  780

Qy    801 IVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVRHEIPGGQLSNL  860
          |:  |:||  || | : | |:| :||: |||| |  :| |||:: |  |:|:||| ::|
Db    781 LYYGLVNGNRQTNLDAQNSQIINHYWEDVRHYYKDFDNALNSPQTEYYIHEMPGGQYTNL  840

Qy    861 RAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAAD  920
          : |  | |  |||: :|| :| :|:| ||  || || ||||: |  ||||||| |||:|
Db    841 QQQAIAVGLGDRWDEVKEMYTVNQMFGDIVKVTPSSKVVGDLIALFMVQNELSEEDVYEK  900

Qy    921 PQKYDIPDSVIAFLRGELGNPPGGWPPLRTRALEGRSEGKAPLTEVPEEQAHLDADDS  980
           ||  |  |:| | | |||||:||  |||:  | ||:|   | |  |::: || |||:
Db    901 GDTIDFPDSVIEFMGEIGQPYGGFPEKLQKLVLKGRT---PLTDRPGALMEPVNFVDV  956

Qy    981 K--------ERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLI  1029
          |        :: |::||  |::| ::: |  | : | || ::|  || :|  |::|:
Db    957 KAELKEKMGYEPTEKDVISYLIYPKVFLDYQDMINKYGDVIVLDTPTFYKGMRLGETIEV  1016
```

```
Db   1008 YGMTLGEIEVEIERGKT-LIVKLISIGEPQPDATRVVYFELNGQPREVVIKDESIKSSV  1066
          |||||||| |:|| ||| |:| :|:| || | |||| |:  ||||||: |:|:| ||:|
Qy   1079 ATAEKADSSNKGHVAAPFAGVVTVTVAE-GDEVKAGDAVAIIEAMKMEATTASVDGKIE  1137
          :|: |  :|  | |  :: ||: ||:|  |||||||||| | |:|||:| |:|: | |:
Db   1067 QERLKADRTNPSHIAASMPGTVIKVLAEAGTKVNKGDHLMINEAMKMETTVQAPFSGTIK  1126
```

```
Qy   1138 RVVPAATKVEGGDLIVVV  1156
          :|  |  :: /|:||| :
Db   1127 QVHVKNGEPIQFGDLLLEI  1145
```

```
RESULT 13
ABB47612
ID   ABB47612 standard; Protein; 1146 AA.
XX
AC   ABB47612;
XX
DT   05-FEB-2002  (first entry)
XX
DE   Listeria monocytogenes protein #316.
XX
KW   Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW   vitamin B12; bacterial infection; disease.
XX
OS   Listeria monocytogenes.
XX
PN   WO200177335-A2.
XX
PD   18-OCT-2001.
XX
PF   11-APR-2001; 2001WO-FR01118.
XX
PR   11-APR-2000; 2000FR-0004629.
XX
PA   (INSP ) INST PASTEUR.
XX
PI   Buchrieser C,  Frangeul L,  Couve E,  Rusniok C,  Fsihi H,  Dehoux P;
PI   Dussurget O,  Chetouani F,  Nedjari H,  Glaser P,  Kunst F,  Cossart P;
PI   Daniels J,  Goebel W,  Kreft J,  Kuhn M,  Ng E,  Vazquez-Boland JA;
PI   Dominguez-Bernal G,  Garrido-Garcia P,  Tierrez-Martinez A,  Amend A;
PI   Chakraborty T,  Domann E,  Hain T,  Berche P,  Charbit A,  Durant L;
PI   Perez-Diaz J,  Baquero F,  Garcia Del Portillo F,  Gomez-Lopez N;
PI   Maduenio E,  De Pablos B,  Wehland J,  Kaerst U,  Entian K,  Hauf J;
PI   Rose M,  Voss H;
XX
DR   WPI; 2002-010914/01.
XX
PT   Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT   and prevention of Listeria and related bacterial infections, and
PT   related polypeptides
XX
PS   Claim 6; SEQ ID No 317; 192pp; French.
XX
CC   The present invention relates to the genome sequence of Listeria
CC   monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC   it are useful for selecting probes and primers for detecting genes in L.
CC   monocytogenes and related organisms, and for studying genetic
CC   polymorphisms and other genomes. The present sequence is a protein
CC   encoded by the genome sequence are useful for raising specific
CC   expressed from the genome sequence of the present invention. Proteins
CC   antibodies, identification of L. monocytogenes and related organisms, and
CC   for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC   B12. The genome sequence and proteins encoded by it are also useful for
CC   selecting compounds that regulate gene expression and cell replication
CC   and modulate L. monocytogenes-related diseases. In addition, the genome
CC   sequence and proteins encoded by it are useful in pharmaceutical and
CC   vaccines compositions for the treatment or prevention of infections by L.
CC   monocytogenes and related organisms.
CC   Note:The sequence data for this patent did not form part of the printed
CC   specification,but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence  1146 AA;
```

```
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence  1147 AA;

Query Match          42.4%;  Score 2489;  DB 22;  Length 1147;
Best Local Similarity  45.2%;  Pred. No. 2.9e-176;
Matches 527;  Conservative 193;  Mismatches 383;  Indels 62;  Gaps 14;
```

```
Qy   30 KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSASEAVRIGTEGSPVKAYL   89
            |::|:|||||||| |||||    :|:||||||     : ||:  |  :| |:|:|| ||
Db    4 KKLLVANRGEIAIRIFRAAAELDISTVAIYSNEDKSSLHRYKADESYLVGSDLGPAESYL   63

Qy   90 DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT  149
            :: ||  : |:  || |||||||||||| ||| ||| ||:|||   | |:  ||| ||:
Db   64 NIERIIDVAKQANVDAIHPGYGFLSENEQFARRCAEEGIKFIGPHLEHLDMFGDKVKART  123

Qy  150 AAKKAGLPVL-AESTPSKNIDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDELRKLA    208
            :| ||| ||  | : | |  ::|::| | |:|||:|| :||| ||| ||  : ||  |
Db  124 TAIKADLPVIPGTDGPIKSYELAKEFAEEAGFPLMIKATSGGGKGMRIVREESELEDAF   183

Qy  209 TEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEI  268
            | |:||:|: ||:  ||:ER ::NP HIEVQ::GD  G::VHL:ERDCS:QRRHQKVVE:
Db  184 HRAKSEAEKSFGNSEVYIERYIDNPKHIEVQVIGDEHGNIVHLFERDCSVQRRHQKVVEV  243

Qy  269 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLV--DEKGNHVFIEMNPRIQVEHTV  326
            ||:  | P LR RIC:DA: ::  I Y  AGTVEFLV  :E:   :FFIE:NPR:QVEHT:
Db  244 APSVGLSPTLRQRICDAAIQLMENIKYVNAGTVEFLVSGDE---FFFIEVNPRVQVEHTI  300

Qy  327 TEEVTEVDLVKAQMRLAAGATL--KELGLTQDK-IKTHGAALQCRITTEDPNNGFRPDTG  383
            TE :T :D:VK Q: :AAGA L   E:     K I T G:A:QCRITTEDP N F PDTG
Db  301 TEMVTGIDIVKTQILVAAGADLFGEEINMPQQKDITTLGYAIQCRITTEDPLNDFMPDTG  360

Qy  384 TITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTV  442
            TI:AYRS GG GVRLD G: Q  GEI:  :DS:LVK:   ::  : A    :R:L E: :
Db  361 TIIAYRSSGGFGVRLDAGDFQGAEISPYIDSLLVKLSTHAISFKQAEEKMVRSLREMRI   420

Qy  443 SGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKP  502
            :GV TNI FL   ::++  T: R:    I   P L :  P :D : ::L:Y: :VT:N
Db  421 RGVKTNIPFLINVMKNKKFTSGDYTTKFIEETPELFDIQPSLDRGTKTLEYIGNVTIN--  478

Qy  503 HGVRPKDVAAPIDKLPNIKDLPLP----------------RGSRDRLKQLGPAAFARD    544
                                            :| :|| : :: GP: :A :
Db  479 --------------GFPNVEKRPKPDYELASIPTVSSSKIASFSGTKQLLDEVGPKGVAEW  525

Qy  545 LREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATVDVAM  604
            :: ::: :TDTTFRDAHQSLLATRVR:  : : : : :::   : : :E WGGAT DVA
Db  526 VKKQDDVLLDTTFRDAHQSLLATRVRTKDMINIASKTADVFKDGFSLEMWGGATFDVAY   585

Qy  605 RFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRI  664
            :FL E PW:RL::LR:A:PNV :QML :  N VGY  YPD V ::FV:E:A :G:D:FRI
Db  586 NFLKENPWERLERLRKAIPNVLFQMLRASNAVGYKNYPDNVIHKFVQESAKAGIDVFRI   645

Qy  665 FDALNDVSQMRPAIDAVLEINTAVAVEMAYSGDLSDP-NEKLYTLDYILKMAEEIVKSG   723
            FD:L:DV QM: A::AV E   ::  E  Y:GD: :P NE:  YTL:Y::K:A:E:  :G
Db  646 FDSLNWVDVQMKVANEAVQEAG-KISEGTICYTGDIINPERSNIYTLEYYVKLAKELEREG  704

Qy  724 AHILAIKDMAGLLRPAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVD   783
            :HILAIKDMAGLL:P A:  L: :L: A DLP:H: HDT:G  L TY : A:AG DA:D
Db  705 FHILAIKDMAGLLKPKAAYELIGELKAAVDLPIHLHTHDTSGNGLLTYKQAIDAGVDIID  764

Qy  784 GASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGP  843
            :A:A :SG:TSQPS :::  A: : :R T ::  ::::  :YW :VR :Y  FES    P
Db  765 TAVASMSGLTSQPSANSLYYALNGFPRHLRTDIEGMESLSHYWSTVRTYYSDFESDIKSP  824

Qy  844 TGRVYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDL  903
            :  :Y:HE:PGGQ SNL  QA : GL :RF: :::   V:E:   ::KVTPSSKVVGD:
Db  825 NTEIYQHEMPGGQYSNLSQQAKSLGLGERFDEVKDMYRVNFLFGDIVKVTPSSKVVGDM   884

Qy  904 ALHLVGAGVDPADFAADPQKVDIPDSVIAFLRGELGNPGGWPEPLRTRALEGRSEGKAP   963
            AL::V   :D:  ::::P  :D    :::F::GE:G P  : :  L::  L :: E
Db  885 ALYMVQNDLDEQSVITDGVKLDFPESVVSFFKGEIGQPVNGFNKDLQAVILKGQE---A   940

Qy  964 LTEVPEEEQAHLDADDSKE----------RRNSLNRLLFPKPTEEFLEHRRRFGNTSAL 1012
```

```
Qy 1030 RLPDVRTPLLVRLDAISEPDKGMRNVVANVNGQIRPMRVRDRSVESVIATAEKADSSNK 1089
            ::|::||||||  :||| ||  :||  :|:|:| | |||:|
Db 1017 ELEKGKI-LLIKLNSIGEPIADGTRVIYFELNGQPREINIQDMNVQSTVIARRKIDTTNP 1075

Qy 1090 GHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVPAATKVE 1148
            :HV :  :G V::V:VA:GD VK GD :::IEAMKME:TI:A VDG::  : :: :||:
Db 1076 EHVGATMTGSVIQVVVVKKGDSVKKGDPLLITEAMKMETTIQAPFDGEVSSIYVSGDTIE 1135

Qy 1149 GGDLIVVVS 1157
            :GDL:: |:
Db 1136 SGDLLIEVN 1144
```

```
RESULT 14
AAU33972
ID   AAU33972 standard; Protein; 1147 AA.
XX
AC   AAU33972;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus cellular proliferation protein #248.
XX
KW   Antisense; prokaryotic cellular proliferation protein;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI   Yamamoto RT, Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   N-PSDB; AAS51831.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Example 3; Seq ID No 5468; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence represents an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
```

```
Query Match              42.0%;    Score 2465;   DB 22;    Length 1151;
Best Local Similarity   45.4%;    Pred. No. 1.8e-174;
Matches 524;  Conservative 198;  Mismatches 399;  Indels 34;  Gaps 14;

Qy    26 LPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPV   85
         : :.:||.|||  | .|:||  | :::|| |||:||.:|| ||. :.|| |.| |:::||
Db     5 LKQIKKLLVANRGEIAIRIFRAAAELNISTVAIYSNEDKSSLHRYKADESYLVGSDLGPA   64

Qy    86 KAYLDIDEIIGAAKKVKADAIYPGVGFLSENAQLARECAENGITFIGPTPEVLLTGDKS  145
         ::|| :||| || |  || ||||| |||||||||  ||| |  |||||| ||:||:||||
Db    65 ESYLNIERIIEVALRAGVDAIHPGYGFLSENEQFARRCAEEGIKFIGPHLEHLDMFGDKV  124

Qy   146 RAVTAAKKAGLPVL-AESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDEL  204
         : :||:| |:|: . : ||| | |.||.|||||||.|| |||:|||||||:|||||||||
Db   125 KARTAINANLPVIPGTDGPIESFEAAEQFANEAGQYPLM.KATSGGGGKGMRIVRESSEL  184

Qy   205 RKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQK  264
         |||: :. :|  .:|: :.|.||:|| |: |:|.|| :.  |: ||.||:|||..||||||
Db   185 EDAFHRAKSEAEKSFGNSEVYIERYIDNPKHIEVQVIGDEFGNIIHLYERDCSVQRRHQK  244

Qy   265 VVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLV--DEKGNHVFIEMNPRIQV  322
         ||:|| :|:  |  :|: :.:: ||::|  |  ||||| |::  ||||| |.|| || ||
Db   245 VVEVAPSVGLSNKLRERICDAAIQLMENIKYVNAGTVEFLVSGDE---FFIEVNPRVQV  301

Qy   323 EHTVTEEVTEVDLVKAQMRLAAGATL--KELGLTQDKIKTHGAALQCRITTEDPNNGFR  379
         |||  :: |::|||| || ||: |:|    ||| | |. |:||  :|| |||||| |:||
Db   302 EHTITEMITGIDIVKTQILVADGESLFGDKISMPQQNEIQTLGYAIQCRITTEDPTNDFM  361

Qy   380 PDTGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALA  438
         || ||||| || ||  |||  :: |: |||   | |:.|| || || |  :|||||: |
Db   362 PDSGTITAYRSSGGFGVRLDAGDGFQGAEISPYIDSLLVKLSTHAVSFKQAEEKMERSLR  421

Qy   439 EFTVSGVATNIGFLRALLREDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVT  498
         | |||| | |||| | |:  :|  |:. :  : |.|||  :. : |::.|  ::.::|: |
Db   422 EMRIRGVKTNIPFLINVMRNDKFRSGDYTTKFIETPELFDIAPTLDRGTKTLEYIGNVT  481

Qy   499 VNKPHGV--RPKDVAAPIDKLPNIKDLPLPR--GSRDRLKQLGPAAFARDLREQDALAVT  554
         :|| | .   | |  :| ||.|| ||: : :|  || |.|||.|| |   :.|:||  |:|
Db   482 INGFPNVEKRKP-EYESTKIPKISQKKINQLFGTKQILEQHGPTGVTNWVREQEDVLIT  540

Qy   555 DTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVPAWGGATYDVAMRFLFEDPWDR  614
         ||||||||||||||||||||| | :|| |.|||||.||.| |||.|| |. :.|| | |
Db   541 DTTFRDAHQSLLATRVRTKDMMNIASKTAEVFKDSFSLEMWGGATFDVAYNFLKENPWER  600

Qy   615 LDELREAMPNVNIQMLLRGRNTVGTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQM  674
         |: |::|||:| :  |||| :| | | |: ::|| || | :. ||:||||| || |.| .
Db   601 LERLRKAIPNVLFQMLLRASNAVGYKNYPDNVIKKFVHESAKAGVDVFRIFDSLNWVDQM  660

Qy   675 RPAIDAVLETNTAVAEVAMAYSGD-LSDPNEKLYTLDYYLKMAEEIVKSGAHIAIKDMA  733
         |. :.:|| . : | :|||||    .:.|| |||||||||||| | .||||.||||||||
Db   661 KVANEAVQEAG-MVSEGTICYTGDILNAERSNIYTIDYYVKMAKFLEREGFHILAIKDMA  719

Qy   734 GLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTT  793
         ||| |.| :|| |  || :  |: ||||:| |.| |||: ::|:.: |:: :|   .| |
Db   720 GLLKPKAAYELIGELREATHLPIHLHTHDTSGNGLLTYKQAIDAGVDIDTAVASMSGLT  779

Qy   794 SQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIP  853
         || |||.|. ||.||| ||| |.. .| . ||: . || ||: | || .||  . ..|:
Db   780 SQPSANSLYYALNGFPRNLRTDIDGLEELSHYWSVVRPYYADFESDIKSPNTEIYQHEMP  839

Qy   854 GGQLSNLRAQATAIGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVD  913
         |||:|| :|: | :|| || |:::. :.|.|::| | .| ||||||.||  | :|| | :
Db   840 GGQYSNLSQQAKSLGLGERFDEVKEMYRRVNFLFGDLVKVTPSSKVVGDMALIMVQNDLD  899

Qy   914 PADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQA  973
         |. : :. :|| ||.|  ||||| |||| |||  |::|| || |:|| | |||.|||.| |
Db   900 EDTVINDGYKLDFPESVVSFFKGDIGQPVNGFNKKLQDVILKQQ---PITERPGEYLE  955

Qy   974 HLDAD--------DSKERRNSLNRLLFPKPTEEFLEHRRFGNTSALDDREFFYGLV  1022
         : :: .         :|:||:| |: |||.||.|| :| |: :|:||||:||:.|
Db   956 PVDFEAIRQELSDIQQDEVTEQDIISYVLYPKVYKQYIQTKEQFGNVSLLDTPTFLFGMR  1015

Db   941 LTARPGEYLEPVDFEKVRELLEEQQGPVTEQDIISVYLYPKVYEQYIQTRNQYGNL.SLL  1000
         || | |||||||||||||| || || ||||||||| |||||||.||||| : .||: |||
Qy  1013 DDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDR  1072

Db  1001 DTPTFFFGMRNGETVEIEI-DKGKRLIIKLEFISEPDENGNRTIYYAMNGQARRIYIKDE  1059
         ::| |||||||||.| |:|  ||.||:||||.|||||:|||||||| ||| |  |:|::|
Qy  1073 SVESVTATAEKADSSNKGHVAAPFAGVVT-VTVAEGDEVKAGDAVAIIEAMKMEATTPAS  1131

Db  1060 NVHTNANVKPKADKSNPSHIGAQMPGSVTEVKKVSVGETVKANQPLLITEAMKMETTIQAP  1119
         |. | |:||| | | | .|.| |:||:||:||.||||||.|| ||||:| ||:|| ||.|
Qy  1132 VDGKIERVVPAATKVEGGDLIVVV  1156

Db  1120 FDGVIKQVTVNNGDTIATGDLLIEI  1144
```

RESULT 15
AAG82677
ID   AAG82677 standard; Protein; 1151 AA.
XX
AC   AAG82677;
XX
DT   03-SEP-2001 (first entry)
XX
DE   S. epidermidis open reading frame protein sequence SEQ ID NO:2448.
XX
KW   Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW   vaccination; endocarditis.
XX
OS   Staphylococcus epidermidis.
XX
PN   WO200134809-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30782.
XX
PR   09-NOV-1999; 99US-0164258.
XX
PA   (GLAX) GLAXO GROUP LTD.
XX
PI   Kimmerly WJ;
XX
DR   WPI; 2001-316495/33.
DR   N-PSDB; AAH53527.
XX
PT   Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT   useful for vaccinating against infections, e.g. endocarditis -
XX
PS   Claim 18; Page 644-645; 2188pp; English.
XX
CC   AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC   (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC   (I) and (II) can have antibacterial activity and therefore can be used
CC   in vaccination. The nucleic acids (I) may be used to produce the
CC   S. epidermidis polypeptides (II) via the production of vectors
CC   containing them which are used to produce hosts cells which express the
CC   polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC   used to vaccinate subjects and to raise antibodies against the bacteria.
CC   The polypeptides may also be used to assay for other inhibitors of their
CC   activity and therefore identify compounds that may be used for the
CC   treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC   AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC   polynucleotide sequences from the present invention. AAH55091 to
CC   AAH55098 represent oligonucleotide sequences and primers which are used
CC   in the exemplification of the present invention.
CC   N.B. The present invention specifically claims all the polynucleotide
CC   sequences given in the sequence listing of the present specification,
CC   however the sequence listing only goes up to SEQ ID NO:4454 so even
CC   though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC   no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ   Sequence 1151 AA;

```
Qy     1023  EGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAE 1082
                   -  : - |  - |   :||||:: |:||||:  :||| | :  ::| :|::
Db     1016  NGETVEIEI-DTGKRLIIKLETISEPDENGKRTIYAMNGQARRIYIQDENVKTNANVKP 1074

Qy     1083  KADSSNKGHVAAPFAGVVT-VTVAEBGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVV 1141
                   |||| |: |: ||||:   |||:: : |:||||:: |||||| ||: |:::|
Db     1075  KADKSNPNHIGAQMPGSVTEVKVSVGDEVQANQPLLITEAMKMETTIQAPFDGIIKQINV 1134

Qy     1142  PAATKVEGGDLIVVV 1156
                  : |||:|:  :
Db     1135  ANGDAIATGDLLVEI 1149
```

Search completed: September 24, 2003, 15:55:21
Job time : 70.4737 secs

```
            Copyright  GenCore version 5.1.6
            (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:     September 24, 2003, 15:27:07 ; Search time 18.1332 Seconds
                                           (without alignments)
                                           2699.668 Million cell updates/sec

Title:      US-09-974-973A-2
Perfect score: 5865
Sequence:   1 MTAITLGGLLLKGIITLVST........RVVVPAATKVEGGDLIVVVS 1157

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:   328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters:   328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_AA:*
  1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
  2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
  3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
  4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
  5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
  6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5759	98.2	1140	3	US-09-220-081-2	Sequence 2, Appli
2	5759	98.2	1140	4	US-09-677-575-2	Sequence 2, Appli
3	3484	59.4	1124	4	US-08-311-731A-10	Sequence 10, Appl
4	2465	42.0	1154	4	US-09-134-001C-3428	Sequence 3428, Ap
5	2182	37.2	973	4	US-09-107-532A-4810	Sequence 4810, Ap
6	1192	20.3	694	4	US-09-433-043B-126	Sequence 126, App
7	1084	18.5	456	4	US-09-634-238-276	Sequence 276, App
8	1039.5	17.7	593	4	US-09-433-043B-122	Sequence 122, App
9	1032.5	17.6	447	1	US-08-611-107-6	Sequence 6, Appli
10	1032.5	17.6	447	2	US-08-422-560A-6	Sequence 6, Appli
11	1032.5	17.6	447	3	US-08-468-793-6	Sequence 6, Appli
12	1005.5	17.1	453	1	US-08-611-107-8	Sequence 8, Appli
13	1005.5	17.1	453	2	US-08-422-560A-8	Sequence 8, Appli
14	1005.5	17.1	453	3	US-08-468-793-8	Sequence 8, Appli
15	1003.5	17.1	453	4	US-09-433-043B-121	Sequence 121, App
16	996.5	17.0	453	1	US-07-956-700B-6	Sequence 6, Appli
17	996.5	17.0	453	2	US-08-476-537-6	Sequence 6, Appli
18	996.5	17.0	453	3	US-08-485-607-6	Sequence 6, Appli
19	996.5	17.0	453	2	US-08-475-879-6	Sequence 6, Appli
20	996.5	17.0	453	4	US-09-433-043B-6	Sequence 6, Appli
21	973	16.6	474	4	US-09-328-352-7562	Sequence 7562, Ap
22	964.5	16.4	1116	4	US-09-252-991A-24374	Sequence 24374, A
23	954	16.3	605	4	US-09-433-043B-123	Sequence 123, App
24	940	16.0	448	5	US-08-074-121-3	Sequence 3, Appli
25	940	16.0	448	5	PCT-US94-06447-3	Sequence 3, Appli
26	937.5	16.0	454	4	US-09-198-452A-197	Sequence 197, App
27	937.5	16.0	701	4	US-09-252-991A-27999	Sequence 27999, A
28	933	15.9	427	1	US-07-956-700B-3	Sequence 3, Appli
29	933	15.9	427	1	US-08-476-537-3	Sequence 3, Appli
30	933	15.9	427	1	US-08-485-607-3	Sequence 3, Appli
31	933	15.9	427	2	US-08-475-879-3	Sequence 3, Appli
32	933	15.8	427	4	US-09-433-043B-3	Sequence 3, Appli
33	928	15.8	536	3	US-08-662-344-2	Sequence 2, Appli
34	921.5	15.7	465	4	US-09-252-991A-26980	Sequence 26980, A
35	921	15.7	449	1	US-08-074-121-6	Sequence 6, Appli
36	921	15.7	449	5	PCT-US94-06447-6	Sequence 6, Appli
37	910	15.5	449	4	US-09-252-991A-26143	Sequence 26143, A
38	902	15.4	676	4	US-09-328-352-6725	Sequence 6725, Ap
39	892	15.2	652	4	US-09-328-352-5587	Sequence 5587, Ap
40	888	15.1	670	4	US-09-328-352-6420	Sequence 6420, Ap
41	869	14.8	573	4	US-09-252-991A-26616	Sequence 26616, A
42	849	14.5	1212	4	US-09-134-001C-3604	Sequence 3604, Ap
43	843	14.4	461	4	US-09-252-991A-19134	Sequence 19134, A
44	842.5	14.4	612	4	US-09-433-043B-125	Sequence 125, App
45	832.5	14.2	453	4	US-09-252-991A-19829	Sequence 19829, A

ALIGNMENTS

```
RESULT 1
US-09-220-081-2
;  Sequence 2, Application US/09220081
;  Patent No. 6171833
;  GENERAL INFORMATION:
;  APPLICANT: Sinskey, Anthony J.
;  APPLICANT: Lessard, Philip A.
;  APPLICANT: Willis, Laura B.
;  APPLICANT: Stephanopoulos, Gregory
;  TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
;  FILE REFERENCE: 1533.0790000
;  CURRENT APPLICATION NUMBER: US/09/220,081
;  CURRENT FILING DATE: 1998-12-23
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 2
;  LENGTH: 1140
;  TYPE: PRT
;  ORGANISM: Corynebacterium glutamicum
US-09-220-081-2

Query Match             98.2%;   Score 5759;   DB 3;   Length 1140;
Best Local Similarity   99.4%;
Matches 1133;   Conservative    5;   Mismatches    2;   Indels    0;   Gaps    0;

Qy     18  VSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSPHRSFASEAVR   77
           :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSPHRSFASEAVR   60

Qy     78  IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  137
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  120

Qy    138  LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRF  197
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF  180

Qy    198  VSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCS  257
           :||:||||||||||||||||||||||:||||||||||||||||||||:||||||||||||
Db    181  VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS  240

Qy    258  LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN  317
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN  300

Qy    318  PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG  377
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG  360

Qy    378  FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  437
```

US-09-677-575-2

Query Match 98.2%;   Score 5759;   DB 4;   Length 1140;
Best Local Similarity 99.4%;   Pred. No. 0;
Matches 1133;   Conservative 5;   Mismatches 5;   Indels 2;   Gaps 0;

```
Qy   18 VSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR   77
          :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR   60

Qy   78 IGTEGSPVKAYLDIEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  137
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGTEGSPVKAYLDIEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  120

Qy  138 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRF  197
         |||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db  121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRF  180

Qy  198 VSSPDELRKLATASREAEAAFGDGSVYYERAVINPQHIEVQILGDRTGEVVHLYERDCS  257
         |:|||||||||:||||||||||||.||.||||||||||||||||||:|||||||||||
Db  181 VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS  240

Qy  258 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN  317
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN  300

Qy  318 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG  377
         ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db  301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATIKELGLTQDKIKTHGAALQCRITTEDPNNG  360

Qy  378 FRPDTGTTTAYRSPGGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  437
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 FRPDTGTTTAYRSPGGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  420

Qy  438 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADV  497
         |||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db  421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV  480

Qy  498 TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  557
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  540

Qy  558 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  617
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  600

Qy  618 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  677
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  660

Qy  678 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR  737
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR  720

Qy  738 PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  797
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  780

Qy  798 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL  857
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL  840

Qy  858 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF  917
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF  900

Qy  918 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA  977
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA  960

Qy  978 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1037
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020
```

```
Db  361 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        FRPDTGTTTAYRSPGGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  420

Qy  438 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADV  497
        |||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db  421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV  480

Qy  498 TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  557
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  540

Qy  558 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  617
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  600

Qy  618 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  677
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  660

Qy  678 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR  737
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR  720

Qy  738 PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  797
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  780

Qy  798 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL  857
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL  840

Qy  858 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF  917
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF  900

Qy  918 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA  977
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA  960

Qy  978 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1037
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020

Qy 1038 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1097
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080

Qy 1098 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATTTASVDGKIERVVVPAATKVEGGDLIVVVS 1157
        |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db 1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATTTASVDGKIDRVVVPAATKVEGGDLIVVVS 1140
```

RESULT 2
US-09-677-575-2
; Sequence 2, Application US/09677575
; Patent No. 6403351
; GENERAL INFORMATION:
; APPLICANT: Sinskey, Anthony J.
; APPLICANT: Lessard, Philip A.
; APPLICANT: Willis, Laura B.
; APPLICANT: Stephanopoulos, Gregory
; TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
; FILE REFERENCE: 1533.0790000
; CURRENT APPLICATION NUMBER: US/09/677,575
; CURRENT FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 09/220,081
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1140
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum

Db   182 EAASREAESAFGDPTVYLEQAVINPRHIEVQILADNLGDVIHLYERDCSVQRRHQKVIEL 241
Qy   269 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE 328
Db   242 APAPHLDAELRYKMCVDAVAFARHIGYSCAGTVEFLLDERGEYVFIEMNPRVQVEHTVTE 301
Qy   329 EVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAY 388
Db   302 EITDVDLIVASQLRIAAGETLEQLGLRQEDIAPHGAALQCRITTEDPANGFRP-TRAGSAR 360
Qy   389 RSPGGAGVRLDGAAQLGGEITAHFDSMLVKM-----TCRGSDFETAVARAQRALAEFTVS 443
Db   361 CDPPAVPVSAWTAAP------TWRRNQPVLRLHAGQADLSGRDLPTAVSRARRAIAEFRIR 415
Qy   444 GVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKPH 503
Db   416 GVSTNIPFLQAVLDDPDFRAGRVTTSFIDERPQLLTARASADRGTKILNFLADVTVNNPY 475
Qy   504 GVRPKDVAAPIDKLPNIKDLPL----PRGSRDRLKQLGPAAFARDLREQDALAVTDTFR 559
Db   476 GSRPSTI-YPDDKLP---DLDLRAAPPAGSKQRLVKLGPEGFARWLRESAAVGVTDTTFR 531
Qy   560 DAHQSLLATRVRSFALKPAARAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELR 619
Db   532 DAHQSLLATRVRTSGLSRVAPYLARTMPQLLSVECWGGATYDVALRFLKEDPWERLATLR 591
Qy   620 EAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAID 679
Db   592 AAMPNICLQMLLRGRNIVGYTPYPEIVTSAFVQEATATGIDIFRIFDALNIESMRPAID 651
Qy   680 AVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYILKMAEEIVKSGAHIIAIKDMAGLLRPA 739
Db   652 AVRETGSAIAEVAMCYTGDLITDPGEQLYTLDYYLKLAEQIVDAGAHVLAIKDMAGLLRPP 711
Qy   740 AVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLS 799
Db   712 AAQRLVSALRSRFDLPVHLHTHDTPGGQLASVYAAWHAGADAVDGAAAPLAGTTSQPALS 771
Qy   800 AIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQLSN 859
Db   772 SIVAAAHTEYDTGLSLSAVCALEPYWEALRKVYAPFESGLPGTGRVYHHEIPGGQLSN 831
Qy   860 LRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAA 919
Db   832 LRQQAIALGLGDRFEEIEEAYAGADRVLGRLVKVTPTSKVVGDLALALVGAGVSADEFAS 891
Qy   920 DPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDADD 979
Db   892 DPARFGIPESVLGFLRGELGDPPGGWPEPLRTAALAGRGAAR-PTAQLAADDEIALSSVG 950
Qy   980 SKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLL 1039
Db   951 AK-RQATLNRLLFPSPIKEFNEHREAYGDTSQLSANQFFYGLRQGEEHRVKL-ERGVELL 1008
Qy 1040 VRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPPAGV 1099
Db 1009 IGLEAISEPDERGMRTVMCILNGQLRPVLVRDRSIASAVPAAEKADRGNPGHIAAPPAGV 1068
Qy 1100 VTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVV 1141
Db 1069 VTIVGVCGERVGAGQTIATIEAMKMEAPITAPVAGTVERVAV 1110

RESULT 4
US-09-134-001C-3428
; Sequence 3428, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C

Qy 1038 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPPA 1097
Db 1021 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPPA 1080
Qy 1098 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVPAATKVEGGDLIVVVS 1157
Db 1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS 1140

RESULT 3
US-08-311-731A-10
; Sequence 10, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
;   STREET: 600 ATLANTIC AVENUE
;   CITY: BOSTON
;   STATE: MASSACHUSETTS
;   COUNTRY: USA
;   ZIP: 02210
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/311,731A
;   FILING DATE:
;   CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
;   NAME: GATES, EDWARD R.
;   REGISTRATION NUMBER: 31,616
;   REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 617/720-3500
;   TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1124 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
;   ORGANISM: MYCOBACTERIUM TUBERCULOSIS
US-08-311-731A-10

Query Match   59.4%; Score 3484; DB 4; Length 1124;
Best Local Similarity   62.2%; Pred. No. 4.6e-267;
Matches 698; Conservative 151; Mismatches 251; Indels 22; Gaps 9;

Qy   29 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY   88
Db   2 FSKVLVANRGEIARAFRAAYELGVGTVAVYPYEDRNSQHRLKADESYQIGDIGHPVHAY   61
Qy   89 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 148
Db   62 LSVDEIVATARRAGADAIYPGYGFLSENPDLAAACAAAGISFVGPSAEVLELAGNKSRAI 121
Qy   149 TAARKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKLA 208
Db   122 AAREAGLPVLMSSAPSASVDELLSVAAGMPFPLFVKAVAGGGGRGMRRVGDIAALPEAI 181
Qy   209 TEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVHLYERDCSLQRRHQKVVEI 268

```
;  CURRENT FILING DATE: 1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
;  SEQ ID NO 3428
;  TYPE: PRT
;  ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3428

Query Match      42.0%;   Score 2465;   DB 4;   Length 1154;
Best Local Similarity  45.4%;   Pred. No. 2.7e-186;
Matches  524;   Conservative 198;   Mismatches 399;   Indels 34;   Gaps    14;

Qy    26 LPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPV 85
            ||||  ::|||||||||  :|||||| | :: |||   :    |  |: :|||  |:||
Db     6 LKQIKKLLVANRGEIAIRIFRAAAELNISTVAIYSNEDKSSLHRYKADESYLVGSDLGPA 65

Qy    86 KAYLDIDEILIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKS 145
            :::|:|:|: | |::||||||||||||||||::|||.|| :||||||| || |.: |||
Db    66 ESYLNIERIIEVALRAGVDAIHPGYGFLSENEQFARRCAEEGIKFIGPHLEHLDMFGDKV 125

Qy   146 RAVTAAKKAGLPVL-AESTPSKNIDDIYKSAEGCTYPIFVKAVAGGGRGMRFVSSPDEL 204
            :|  |   .||||| ::  :|:  | || ||| |||| :|||:|||| |||.||  :||
Db   126 KARTTAINANLPVIPGTDGPIESFEAAEQFANEAGYPLMIKATSGGGKGMRIVRESSEL 185

Qy   205 RKLATEASREAAFAGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQK 264
            .: |  |:.||  |.| ::|:||. ::||:||||:.||  |:::||||||||:|||||||
Db   186 EDAFHRAKSEAEKSFGNSEVYIIERYIDNPKHIEVQVIGDEFGNIIHLYERDCSVQRRHQK 245

Qy   265 VVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLV--DEKGNHVFIEMNPRIQV 322
            |||:||:.  .:.||:|||.||:|  : :.| | ||||||| |  :  :|::|:|||:||
Db   246 VVEVAPSVGLSNKLRERICDAAIQLMENIKYVNAGTVEFLVSGDE--FFFIEVNPRVQV 302

Qy   323 EHTVTEEVTEVDLVKAQMRLAAGATL--KELGLTQDKIKTHGAALQCRITTEDPNNGFR 379
            |||:||  | :|:||:|: |  |.:|   |:  :  ||||  ||:|||||||||| | |:
Db   303 EHTITEMITGIDIVKTQILVANGESLFGDKISMPQQNEIQTLGYAIQCRITTEDPTNDFM 362

Qy   380 PDTGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALA 438
            ||:|||:|||| || |||||  |  |.|||::||||:|||::    :.|: |  |:|:|
Db   363 PDSGTIIAYRSSGGFGVRLDAGDGFQGAEISPYYDSLLVKLSTHAVSFKQAEEKMERSLR 422

Qy   439 EFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVT 498
            |  : ||.|||.||  ::|::||:|: ||| | |:  :||  |  |::|:|: ||:.:||
Db   423 EMRIRGVKTNIPFLINVMRNDKFRSGDYTTKFIEETPELFDIAPTLDRGTKTLEYIGNVT 482

Qy   499 VNKPHGV--RPKDVAAPIDKLPNIKDLPLPR-GSRDRLKQLGPAAFARDLREQDALAVT 554
            :| |  |   ||| :  || |.|:| :  |:   : :: ||| ||    :.| ||:|:|
Db   483 INGFPNVEKRPKP-EYESTKIPKISQKKINQLFGTKQILEQHGPTGVTNWVREQEDVLIT 541

Qy   555 DTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDR 614
            |||||||||||||||||: |:  :|.|.||: |::|:||:|||||:||| .||:|:||:|
Db   542 DTTFRDAHQSLLATRVRTKDMMNIASKTAEVFKDSFSLEMWGGATFDVAYNFLKENPWER 601

Qy   615 LDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQM 674
            |: ||:|:|||:::|||| .|:||| .||.:|  .||.|:|.:|||:  |||:||:|.|||
Db   602 LERLRKAIPNVLFQMLLRASNAVGYKNYPDNVIKKFVHESAKAGVDVPFIFDSLNWVDQM 661

Qy   675 RPAIDAVLETNTAVAEVAMAYSGD-LSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMA 733
            : ||:||:| ||:| | :: | || |::.|.::||||||:|||:|: :.| ||||||||||
Db   662 KVANEAVQEAG-MVSEGTICYTGDILNAERSNIYTLDYYVKMAKELEREGFHILAIKDMA 720

Qy   734 GLLRPAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTT 793
            |||:||||  |: |||||  ||:|:|||||:|: | |: | .||.||.:| |:| |||
Db   721 GLLKPAAVELIGELREATHLPIHLHTHDTSGNGLLTYKQAIDAGVDIIDTAVASMSGLT 780

Qy   794 SQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIP 853
            |||| :: :.| | .|::|.||::||:.   |:|:: | | ||:|:   . | :|:||:|
Db   781 SQPSANSLYYALNGFPRNLRTDIGLEELSHYKNSVVRPYYADFESDIKSPNTEIYQHEMP 840

Qy   854 GGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPIKVTPSSKVVGDLALHLVGAGVD 913
            |||:|||. ||: |||.:||: ::::| .||::  : :|||||||||||:||:::.    :
Db   841 GGQYSNLSQQAKSLGLGERFDEVKEMYRRVNFLFGDLVKVTPSSKVVGDMALYMVQNDLD 900

Qy   914 PADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEQA 973
            . ||  | :  |  : |::|| | :| :    .    |||  .| ::|:||:|:||||:|
Db   901 EDTVINDGYKLDFPESVVSFFKGDIGQPVNGFNKKLQDVILKGQQ---PIIERPGEYLE 956

Qy   974 HLDAD--------DSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDREFYGLV 1022
            ::|.|        |.|.  | :: :| ||: ||| ||::.|||.|.||  | | :
Db   957 PVDFEAIRQELSDIQQDEVTEQDIISVLYPKVYKQYIQTKEQFGNVSLLDTPTLFGMR 1016

Qy  1023 EGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAE 1082
            .| | :  ::| : : |: ||:|||||:.| | :  | ||| |: :  : |:| . ..:
Db  1017 NGETVEIEI-DTGKRLIIKLETISEPDENGKRTIYYAMNGQARRIYIQDENVKTNANVKP 1075

Qy  1083 KADSNKGHVAAPFAGVGT-VTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVV 1141
            |||.:  ::.. :.| :| | |:.||||:|.:   |||||||||:|  . :||||:::|
Db  1076 KADKSNPNHIGAQMPGSVTEVKVSVGDEVQANQPLLITEAMKMETTIQAPFDGIIKQINV 1135

Qy  1142 PAATKVEGGDLIVVV 1156
            . | |  |||::::
Db  1136 ANGDAIATGDLLVEI 1150

RESULT 5
US-09-107-532A-4810
;  Sequence 4810, Application US/09107532A
;  Patent No. 6583275
;  GENERAL INFORMATION:
;  APPLICANT: Lynn A Doucette-Stamm  and David Bush
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;  ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;  NUMBER OF SEQUENCES: 7310
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;  STREET: 100 Beaver Street
;  CITY: Waltham
;  STATE: Massachusetts
;  COUNTRY: USA
;  ZIP: 02354
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE: CD/ROM ISO9660
;  COMPUTER: PC
;  OPERATING SYSTEM: <Unknown>
;  SOFTWARE: ASCII
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/09/107,532A
;  FILING DATE: 30-Jun-1998
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER: 60/085,598
;  FILING DATE: 14 May 1998
;  APPLICATION NUMBER: 60/051571
;  FILING DATE: July 2, 1997
;  ATTORNEY/AGENT INFORMATION:
;  NAME: Ariniello, Pamela Deneke
;  REGISTRATION NUMBER: 40,489
;  REFERENCE/DOCKET NUMBER: GTC-012
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE: (781)893-5007
;  TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4810:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH: 973 amino acids
;  TYPE: amino acid
;  TOPOLOGY: linear
;  MOLECULE TYPE: protein
;  HYPOTHETICAL: YES
;  ORIGINAL SOURCE:
;  ORGANISM: Enterococcus faecium
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (B) LOCATION 1...973
;  SEQUENCE DESCRIPTION: SEQ ID NO: 4810:
```

```
US-09-107-532A-4810

Query Match           37.2%;  Score 2182;  DB 4;  Length 973;
Best Local Similarity 48.1%;  Pred. No. 5.4e-164;
Matches 455;  Conservative 143;  Mismatches 320;  Indels 28;  Gaps 8;

Qy   28  AFKKIIVANRGEIAVRAFRAALETGAATVAIYPREDRGSPHRSFASEAVRIGTEGSPVKA  87
Db   11  SMKKVLVANRGEIAVRVFRACTELGIKTVGIYAEKNEYSVHRFKADEAYLVGQGKKPIDA  70

Qy   88  YLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRA  147
Db   71  YLDIEGIISIAKECGADAIHPGYGLLSENLNFAQRCEEKGIIFVGPKLIHIDIFGDKIKA  130

Qy  148  VTAAKKAGL-PVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRK  206
Db  131  KAARIEAGIASIPGIDGPIASIDDALEFAKQYGYPIMIKAALCGGGRGMRVAHDEKSARE  190

Qy  207  LATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVHLYERDCSLQRRHQKVV   266
Db  191  GYERAKSEAKAAFGSDEVVVEKYIANPKHIEVQILGDTHGNVIHLFERDCSVQRRHQKVV  250

Qy  267  EIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTV  326
Db  251  EVAPCVSMNEQQRQKICQAAVPLMKVYGVYNARTVEFLV-EGDDLYFIEVNPRVQVEHTI  309

Qy  327  TEEVTEVDLVKAQMRLAAGATL-KELGLTQDK-IKTHGAALQCRITTEDPNNGFRPDTGT  384
Db  310  TEMIDIDIVTTQLLIAQGLDLHKEIGLFQQEGIKLNGSAIQCRITTEDPLNNFLPDTGK   369

Qy  385  ITAYRSPGGAGVRLD-GAAOLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVS  443
Db  370  IDTYRSPGGFGVRLDVGNAYAGYVVTPYFDSLLVKVCTHGATFETAIQKMERCLKEFRIR  429

Qy  444  GVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPADDEQGRILDYLADVTVNKPH   503
Db  430  GVKTNIPFMLNVITHPEFQSGNAKTFIDSTATLFEFPRLRDRGNKTMKYIGEITVNGFP   489

Qy  504  GVR-------PKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDAL       551
Db  490  GIESGEKPFYEEPRMPKDLITRTDYVT------AKNVLDAKGADALVEWIKGQENL      539

Qy  552  AVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRLFEDP   611
Db  540  LLTDTTFRDAHQSLLATRVRTKDFKQIARLTGEGLPELFSSEMWGGATFDVAYRFLNEDP  599

Qy  612  WDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDV  671
Db  600  WQRLKIRSLMPNTLLQMLFRGSNAVGYSNYPDNVLVEFVKEAAAQGIDVFRIFDSLNWT   659

Qy  672  SQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKD  731
Db  660  PQMEKSIQAVRDTG-KIAEAAICYTGDINDPSRAKYNVQYYKDMAKELEQLGAHIIAIKD  718

Qy  732  MAGLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSG  791
Db  719  MAGLLKPQAYRLISELKETTDLPIHLHTHDTSGNGIITYSAASKAGVDIVDVAMSAMSG   778

Qy  792  TTSQPSLSAIVAAFAHTRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHE   851
Db  779  NTSQPSMSSLYYALVNGPRLPEITIENAQKLNHYWEDVRMYYKPFENGLNAPETEVYMHE  838

Qy  852  IPGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAG   911
Db  839  MPGGQYSNLQQQAKAVGLGHRWDEIKQMYHTVNLMFGDIVKVTPSSKVVGDMALFMVQND  898

Qy  912  VDPADFAADPQKYIDIPDSVIAFLRGELGNPPGGWPEPLETRALEGR  957
Db  899  LTEEDIYEKGETLSFPESVVTFFQQELGQPVGGFFKKLQKIILLKGR  944
```

```
US-09-433-043B-126

Query Match           20.3%;  Score 1192;  DB 4;  Length 694;
Best Local Similarity 27.9%;  Pred. No. 8.9e-86;
Matches 316;  Conservative 102;  Mismatches 239;  Indels 239;  Gaps 11;

Qy   31  KILVANRGEIAVRAFRAALETGAATVAIYPREDRGSPHRSFASEAVRIGTEG--SPVKAY  88
Db   21  KILVANRGEIPIRIFRTAHELSMQTVAIYSHEDRLSTHKQADEAYLVGEVGQYTPVGAY   80

Qy   89  LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV  148
Db   81  LAIDEIISIAQKHQVDFIHPGYGFLSENSEFADKVVKAGITWIGPAEVIDSVGDKVSAR   140

Qy  149  TAAKKAGLPVLAEST-PSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKL  207
Db  141  NLAAKANVPTVPGTPGPIETVEEALDEVNEYGYPVIIKAAFGGGGRGMRVVREGDDVADA  200

Qy  208  ATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVHLYERDCSLQRRHQKVVE   267
Db  201  FQRATSEARTAGFNGTCFVERFIDKPHIEVQLLADNHGNVHLFERDCSVQRRHQKVVE   260

Qy  268  IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT  327
Db  261  VAPAKTLPREVRDAILTDAVKLIAKECGYRNAGTAEFLVDNQNRHYFLEINPRIQVEHTIT 320

Qy  328  EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPDNNGFRPDTGTITA 387
Db  321  EEITGIDIVAAQIQ-AAGASLPQLGLFQDKITTRGFAIQCRITTEDPAKNFQPDTGRIEV  379

Qy  388  YRSPGGAGVRLDGA-AQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVA  446
Db  380  YRSAGGNGVRLDGNAYAGTIISPHYDSMLVKCSCSGSTYEIVRRKMIRALLEFRIRGVK   439

Qy  447  TNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKPHGVR  506
Db  440  TNIPFLLTLLTNPVFIEGTYWGTFIDDTPQLFQMVSSQNRAQKLLHYLADVA-------  491

Qy  507  PKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDAHQSLL  566
Db  492  ------------------------------------------------------------  491

Qy  567  ATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRLFEDPWDRLDELREAMPNVN   626
Db  492  ------------------------------------------------------------  491

Qy  627  IQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETNT  686
Db  492  ------------------------------------------------------------  491

Qy  687  AVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLVT  746
Db  492  ------------------------------------------------------------  491

RESULT 6
US-09-433-043B-126
```

```
Db   492  -------------------------------------------------              491
Qy   747  ALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFA  806
Db   492  -------------------------------------------------              491
Qy   807  HTRRDTGLSLEAVSDLEPYWEAVRGLYLFPESGTPGTGRVYRHEIPGGGQLSNLRAQATA  866
Db   492  -------------------------------------------------              491
Qy   867  LGLADRFELIEDNYAAVNEMLGRP-TKVTPSSKVVGDLALHLVGAGVDPADFAADPQKYD  925
Db   492  -------------DNGSSIKGQIGLPKLKSNPSV-------------PHSYN           517
Qy   926  IPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDADDSKERRN  985
Db   518  -------------------------------------------------              517
Qy   986  SLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAI 1045
Db   518  -----MYPRVYEDFQKMRETYGDLSVLPTRSFLSPLETDEEIEVVIEQGKT-LIIKLQAV  571
Qy  1046  SEPDDK-GMRNVVANVNGQIRPMRVPDRSVESVTATAEKADSSNKGHVAAPFAGV-VTVT 1103
Db   572  GDLNKKTGEREVYFDLNGEMRKIRVADRSQKVETVTKSKADMHDPLHIGAPMAGVIVEVK  631
Qy  1104  VAEGDEVKAGDAVAIIEAMKMEAIITASVDGKIERVVVPAATKVEGGDLIVVV         1156
Db   632  VHKGSLIKKGQPVAVLSAMKMEMIISSPSDGQVKEVFVSDGENVDSSDLLVLL          684
```

RESULT 7
US-09-634-238-276
; Sequence 276, Application US/09634238
; Patent No. 654772
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
; APPLICANT: Christensson, Anna C.
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; TITLE OF INVENTION: them and methods for using them.
; FILE REFERENCE: 11000.1043U1
; CURRENT APPLICATION NUMBER: US/09/634,238
; CURRENT FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 276
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-634-238-276

Query Match   18.5%;   Score 1084; DB 4;   Length 456;
Best Local Similarity 51.3%;   Pred. No. 1.6e-77;
Matches 230; Conservative 66; Mismatches 142; Indels 10; Gaps 7;

```
Qy  224  VYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRIC  283
Db   12  MYVEKYIASAKHVEVQVLGDEHGHLILHLFERDCSVQRRQKVVEIAPAVALPLALRNRIC   71
Qy  284  ADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLA  343
Db   72  QSAVDLMASLHYENAGTVEFLVD-GDQYYFIEVNPRVQVEHTITFLITGVDIVQSLRIA  130
Qy  344  AGATL-KELGL-TQDKIKTHGAALQCRITTEDPNNGFRPDTGITTAYRSPGGAGVRLD-G 400
```

```
Db   131  AGADLFADLHLPQQDALRENGAAIQCRITTEDPENNFMPDTGTINTYRSPGFGIRLDVG  190
Qy   401  AAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREED 460
Db   191  NAYAGAVVSPYFDSLLVKASVHAPSFPAAVAKMQRALHEFQITGVKTNVAFLEHLLATQT 250
Qy   461  FTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKPHGV-RPKDVAAPIDKLPN 519
Db   251  FRTGEAETAFIDAHPELLQVQAKPDIASRLLWYISDVTVNGFKGVERQSQKYYP--ELQY 308
Qy   520  IKDLPLPRGSRD--RLKQLGPAAFARDLREQDALAVTDTTFRDAHQSLLATPVRSFALK  576
Db   309  TRHFAAAKPQTDIVALLKNEGAQAVTDWVKAHPALLLTDTTFRDAHQSLFATRMRTRDML 368
Qy   577  PAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNT 636
Db   369  TVAEDMGNGLPNLFSMEVWGGATFDVAYRFLNEDPWVRLKKLRAALPHTLLQMLFRGSNA 428
Qy   637  VGTPYPDSVCRAFVKEAASSGVDIFRI                                   664
Db   429  VGYQNYPDNVIKAFINQAANDGVDVFRI                                  456
```

RESULT 8
US-09-433-043B-122
; Sequence 122, Application US/09433043B
; Patent No. 6399342
; GENERAL INFORMATION:
; APPLICANT: HASELKORN, ROBERT
; APPLICANT: GORNICKI, PIOTR
; TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-CoA CARBOXYLASE
; FILE REFERENCE: ARCD:338US
; CURRENT APPLICATION NUMBER: US/09/433,043B
; CURRENT FILING DATE: 1999-10-25
; PRIOR APPLICATION NUMBER: 08/475,879
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 07/956,700
; PRIOR FILING DATE: 1992-10-02
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 122
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-433-043B-122

Query Match   17.7%;   Score 1039.5; DB 4;   Length 593;
Best Local Similarity 44.3%;   Pred. No. 8.4e-74;
Matches 229; Conservative 80; Mismatches 177; Indels 31; Gaps 6;

```
Qy   29  FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY   88
Db    3  FDKILIANRGEIAILRILRACEMGIATIAVHSTVDRNALHVQLADEAVCIG-EPASAKSY   61
Qy   89  LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV  148
Db   62  LNIPNIIAAALTRNASAIHPGYGFLSENAKFAEICADHHIAFIGPTPEAIRLMGDKSTAK  121
Qy  149  TAAKKAGLPVLAES-----TPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDE 203
Db  122  ETMQKAGVPTVPGSEGLVETEQEGLE---LAKDIGYPVMIKATAGGGGRGMRLVRSPDE  177
Qy  204  LRKLATEASREAEAAFGDGSYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQ  263
Db  178  FVKLFLAAQGEAGAAFGNAGVYIEKFIERPRHIEFQILADNYNGNVIHLGERDCSIQRRNQ 237
Qy  264  KVVEIAPAQHLDPELRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMNPRIQVE  323
Db  238  KLLEEAPSPALDSDLREKMGQAAVKAQFINYAGAGTIEFLLDRSGQFYFEMNNTRIQVE  297
```

RESULT 10 (right column alignment):

```
Qy   29  FKKIIVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY  88
         |||:||||||||| ||: |:||: | : |||::| ::: |:  :|:|
Db    3  FDKILIANRGEIALRILRACEEMGIATIAVHSTVDRNALHVQLADEAVCIG-EPASAKSY  61

Qy   89  LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 148
         |:: ||||  ||| ||||:||||||||| : ||| |:||:||||||||| | ||||: |
Db   62  LNIPNIIAAALTRNASAIHPGYGFLSENAKFAEICADHHIAFIGPTPEAIRLMGDKSTAK 121

Qy  149  TAAKKAGLPVLAES----TPSKNIDDIVKSAEGGTYPIFVKAVAGGGGRGMRFVSSPDE  203
         : ||:|| | | ::    ::|:|| :  ::| |:|| |||:|:|||||||||| |:||||
Db  122  ETMQKAGVPTVPGSEGLVETEQEGLE---LAKDIGYPVMIKATAGGGGRGMRLVRSPDE  177

Qy  204  LRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQ 263
         : || : | : |:||| | :|| |: : ||:|||::||| | |||:|| ||||||:||::|
Db  178  FVKLFLAAQGEAGAAFGNGAGVYIEKFIERPRHIEFQILADNYGNVIHLGERDCSIQRRNQ 237

Qy  264  KVVEIAPAQHLDPELRDRICADAVKFCRSIGYQQAGTVEFLVDEKGNHVFIEMNPRIQVE 323
         | :| || ::||:::  ::: :|||||:| :|| || |||::| |::|:|:|:||| ||||
Db  238  KLLEEAPSPALDSDLREKMGQAAVKAAQFINYTGAGTIEFLLDRSGQFYFMEMNTRIQVE 297

Qy  324  HTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTG 383
         | |||:| ||||: || :: :||  | |||| | : ||| |::|||:|||:||:|||| |
Db  298  HPVTEMVTGVDLLVEQIRIAQGERLR--LTQDQVVLRGHAIECRINAEDPDHDFRPAPG 354

Qy  384  TITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVS 443
          |: |  ||| ||::|     ::  | ||: :: :  ::| ::|  : || ||| | :|:
Db  355  RISGYLPPGGPGVRIDSHVYTDYQIPPYYDSLIGKLIVWGPDRATAINRMKRALRECAIT 414

Qy  444  GVATNIGFLRALLREEDFTSKRIATGFI                                471
         |  | ||| :|: :|| ||:| ||  |:
Db  415  GLPTTIGFHQRIMENPQFLQGNVSTSFV                                442
```

RESULT 10
US-08-422-560A-6
; Sequence 6, Application US/08422560A
; Patent No. 5910626
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert
; APPLICANT: Gornicki, Piotr
; TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
; TITLE OF INVENTION: METHODS FOR USE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Arnold, White & Durkee
;   STREET: P.O. Box 4433
;   CITY: Houston
;   STATE: TX
;   COUNTRY: USA
;   ZIP: 77210-4433
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/422,560A
;   FILING DATE: 14-APR-1995
;   CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/956,700
;   FILING DATE: 02-OCT-1992
; ATTORNEY/AGENT INFORMATION:
;   NAME: Wilson, Mark B.
;   REGISTRATION NUMBER: 37,259
;   REFERENCE/DOCKET NUMBER: ARCD:152/WlM
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 512-418-3000
;   TELEFAX: 512-474-7577
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 447 amino acids
;   TYPE: amino acid

RESULT 9 (left column alignment):

```
Qy  324  HTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTG 383
         | |||:| ||||: || :: :||  | |||| | : ||| |::|||:|||:||:|||| |
Db  298  HPVTEMVTGVDLLVEQIRIAQGERLR--LTQDQVVLRGHAIECRINAEDPDHDFRPAPG 354

Qy  384  TITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVS 443
          |: |  ||| ||::|     ::  | ||: :: :  ::| ::|  : || ||| | :|:
Db  355  RISGYLPPGGPGVRIDSHVYTDYQIPPYYDSLIGKLIVWGPDRATAINRMKRALRECAIT 414

Qy  444  GVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLA--------495
         |  | ||| :|: :|| ||:| ||  |:
Db  415  GLPTTIGFHQRIMENPQFLQGNVSTSFVQE---MNKPLDFNEIRQLLTTIAQTDIAEVT 470

Qy  496  --------DVTVNKPHGVRPKDVAAPIDKLPNIKDLPLP                      526
                 : ::
Db  471  LKSDDFELTVRKAVGVNNSVVPVVTAPLSGVVGSGLP                        507
```

RESULT 9
US-08-611-107-6
; Sequence 6, Application US/08611107
; Patent No. 5801233
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert
; APPLICANT: Gornicki, Piotr
; TITLE OF INVENTION: NUCLEIC ACID COMPOSITIONS ENCODING
; TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE AND USES
; TITLE OF INVENTION: THEREFOR
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Arnold, White & Durkee
;   STREET: P.O. Box 4433
;   CITY: Houston
;   STATE: Texas
;   COUNTRY: United States of America
;   ZIP: 77210
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/611,107
;   FILING DATE: Concurrently Herewith
;   CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US SN 07/956,700
;   FILING DATE: 02-OCT-1992
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US PCT/US93/09340
;   FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US SN 08/422,560
;   FILING DATE: 14-APR-1995
; ATTORNEY/AGENT INFORMATION:
;   NAME: Kitchell, Barbara S.
;   REGISTRATION NUMBER: 33,928
;   REFERENCE/DOCKET NUMBER: ARCD:221
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (512) 418-3000
;   TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 447 amino acids
;   TYPE: amino acid
;   STRANDEDNESS:
;   TOPOLOGY: linear
US-08-611-107-6

Query Match   17.6%;   Score 1032.5;   DB 1;   Length 447;
Best Local Similarity 48.2%;   Pred. No.1.9e-73;
Matches 216;   Conservative 71;   Mismatches 148;   Indels 13;   Gaps 4;

```
;  STRANDEDNESS:
;  TOPOLOGY: linear
US-08-422-560A-6

Query Match       17.6%;  Score 1032.5; DB 2;  Length 447;
Best Local Similarity  48.2%;  Pred. No. 1.9e-73;
Matches 216;  Conservative  71;  Mismatches 148;  Indels  13;  Gaps  4;

Qy   29 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 88
        | ||:|||||||| | :|::|::|: | ||||||| ||||::| |:|:| :|:|
Db    3 FDKILIANRGEIALRILRACEEMGIATIAVHSTVDRNALHVQLADEAVCIG-EPASAKSY 61

Qy   89 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 148
        |:| :||||| :|||:|:||||||||||| |||:| : :||||||||:::|:||||:||
Db   62 LNIPNIIAAALTRNASAIHPGYGFLSENAKFAEICADHHIAFIGPTPEAIRLMGDKSTAK 121

Qy  149 TAAKKAGLPVLAES-----TPSKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDE 203
        :|:||:|||| ||      |:|||:| :|:||:||||:|::||:|||||||:|:|||||
Db  122 ETMQKAGVPTVPGSEGLVETEQEGLE---LAKDIGYPVMIKATAGGGRGMRLVRSPDE 177

Qy  204 LRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQ 263
        |:||: | :|||:|||| ||||:|::||| ::|:|||: ||||:||||||||||:||:
Db  178 FVKLFLAAQGEAGAAFGNAGVYIEKFIERPRHIEFQILADNYGNVIHLGERDCSIQRRNQ 237

Qy  264 KVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVE 323
        |::|:||: || | |:| | ||||| :||:|||||||:||| :|:|:||:||||||||||
Db  238 KLLEEAPSPALDSDLREKMGQAAVKAAQFINYTGAGTIEFLLDRSGQFYFEMEMNTRIQVE 297

Qy  324 HTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTG 383
        | ||| ||||||::|:| | |:|| || || || |||| || |||| ||::| ||| :|
Db  298 HPVTEMVTGVDLLVEQIRIAQGERLR--LTQDQVVLRGHAIECRINAEDPHDFRPAPG 354

Qy  384 TITAYRSPGGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVS 443
        ::| ||| || || |:|| ||| || ||:|||: |  |:| || | |||:||:| ||:
Db  355 RISGYLPPGGPGVRIDSHVYTDYQIPPYYDSLIGKLIVWGPDRATAINRMKRALRECAIT 414

Qy  444 GVATNIGFLRALLREEDFTSKRIATGFI 471
        |::| ||| : |:||| ::|:::| | |
Db  415 GLPTTIGFHQRIMENPQFLQGNVSTSFV 442


RESULT 11
US-08-468-793-6
; Sequence 6, Application US/08468793
; Patent No. 6177267
;  GENERAL INFORMATION:
;  APPLICANT: Haselkorn, Robert
;  APPLICANT: Gornicki, Piotr
;  TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
;  TITLE OF INVENTION: METHODS OF USE
;  NUMBER OF SEQUENCES: 29
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE: Arnold, White & Durkee
;  STREET: P.O. Box 4433
;  CITY: Houston
;  STATE: Texas
;  COUNTRY: United States of America
;  ZIP: 77210
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE: Floppy disk
;  COMPUTER: IBM PC compatible
;  OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
;  SOFTWARE: PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/08/468,793
;  FILING DATE: 06-JUN-1995
;  CLASSIFICATION: 800
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER: US 08/422,560
;  FILING DATE: 14-APR-1995
;  APPLICATION NUMBER: US SN 07/956,700
;  FILING DATE: 02-OCT-1992
;  CLASSIFICATION: 800
```

```
;  APPLICATION NUMBER: PCT/US93/09340
;  FILING DATE: 30-SEP-1993
;  CLASSIFICATION: 800
;  ATTORNEY/AGENT INFORMATION:
;  NAME: Kitchell, Barbara S.
;  REGISTRATION NUMBER: 33,928
;  REFERENCE/DOCKET NUMBER: ARCD:152/KIT
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE: (512) 418-3000
;  TELEFAX: (713) 789-2679
;  TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 6:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH: 447 amino acids
;  TYPE: amino acid
;  STRANDEDNESS: single
;  TOPOLOGY: linear
;  MOLECULE TYPE: protein
US-08-468-793-6

Query Match       17.6%;  Score 1032.5; DB 3;  Length 447;
Best Local Similarity  48.2%;  Pred. No. 1.9e-73;
Matches 216;  Conservative  71;  Mismatches 148;  Indels  13;  Gaps  4;

Qy   29 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 88
        | ||:|||||||| | :|::|::|: | ||||||| ||||::| |:|:| :|:|
Db    3 FDKILIANRGEIALRILRACEEMGIATIAVHSTVDRNALHVQLADEAVCIG-EPASAKSY 61

Qy   89 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 148
        |:| :||||| :|||:|:||||||||||| |||:| : :||||||||:::|:||||:||
Db   62 LNIPNIIAAALTRNASAIHPGYGFLSENAKFAEICADHHIAFIGPTPEAIRLMGDKSTAK 121

Qy  149 TAAKKAGLPVLAES-----TPSKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDE 203
        :|:||:|||| ||      |:|||:| :|:||:||||:|::||:|||||||:|:|||||
Db  122 ETMQKAGVPTVPGSEGLVETEQEGLE---LAKDIGYPVMIKATAGGGRGMRLVRSPDE 177

Qy  204 LRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQ 263
        |:||: | :|||:|||| ||||:|::||| ::|:|||: ||||:||||||||||:||:
Db  178 FVKLFLAAQGEAGAAFGNAGVYIEKFIERPRHIEFQILADNYGNVIHLGERDCSIQRRNQ 237

Qy  264 KVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVE 323
        |::|:||: || | |:| | ||||| :||:|||||||:||| :|:|:||:||||||||||
Db  238 KLLEEAPSPALDSDLREKMGQAAVKAAQFINYTGAGTIEFLLDRSGQFYFEMEMNTRIQVE 297

Qy  324 HTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTG 383
        | ||| ||||||::|:| | |:|| || || || |||| || |||| ||::| ||| :|
Db  298 HPVTEMVTGVDLLVEQIRIAQGERLR--LTQDQVVLRGHAIECRINAEDPHDFRPAPG 354

Qy  384 TITAYRSPGGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVS 443
        ::| ||| || || |:|| ||| || ||:|||: |  |:| || | |||:||:| ||:
Db  355 RISGYLPPGGPGVRIDSHVYTDYQIPPYYDSLIGKLIVWGPDRATAINRMKRALRECAIT 414

Qy  444 GVATNIGFLRALLREEDFTSKRIATGFI 471
        |::| ||| : |:||| ::|:::| | |
Db  415 GLPTTIGFHQRIMENPQFLQGNVSTSFV 442


RESULT 12
US-08-611-107-8
; Sequence 8, Application US/08611107
; Patent No. 5801233
;  GENERAL INFORMATION:
;  APPLICANT: Haselkorn, Robert
;  APPLICANT: Gornicki, Piotr
;  TITLE OF INVENTION: NUCLEIC ACID COMPOSITIONS ENCODING
;  TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE AND USES
;  TITLE OF INVENTION: THEREFOR
;  NUMBER OF SEQUENCES: 40
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE: Arnold, White & Durkee
;  STREET: P.O. Box 4433
;  CITY: Houston
;  STATE: Texas
```

```
Query Match          17.1%;    Score 1005.5;   DB 2;   Length 453;
Best Local Similarity 47.1%;   Pred. No. 2.7e-71;
Matches 209;  Conservative 72;  Mismatches 158;  Indels 5;  Gaps 3;

Qy   29 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 88
         | |::||||||||| | |   ::| :|  ||| ::|| ||:|||| |||: ::  | :|
Db    3 FNKILIANRGEIALRLRTCEELGIGTIAVHSTVDRNALHVQLADEAVCIG-EAASSKSY 61

Qy   89 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 148
         |:| ||| || ::| ||:||||||:|||| |||||:|| |||| ||:|  |||| ||
Db   62 LNIPNIIAAALTRNASAIHPGYGFLAENARFAEICADHHLTFIGPSPDSIRAMGDKSTAK 121

Qy  149 TAAKKAGLPVLAESTP-SKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDELRKL 207
         :   ||:| | :|| | ||| |||  |:|||::| ||| ||||||||||:| | |:|||
Db  122 ETMQRVGVPTIPGSDGLLTDVDSAAKVAAEIGYPVMIKATAGGGRGMRLVREPADLEKL 181

Qy  208 ATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVE 267
         ::  | ||||||| | :|:| :|:||:|||::||| || ||||| |||||:|||||||:|
Db  182 FLAAQGEAEAAFGNPGLYLEKFIDRPRHVEFQILADAYGNVVHLGERDCSIQRRHQKLLE 241

Qy  268 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 327
         :|   |||| || | :|| ||| |:||| |||||||||| |||  |||||:|||||| ||
Db  242 EAPSPALSADLRQKMGDAAVKVAQATGYIGAGTVEFLVDATGNFYFMEMNTRIQVEHPVT 301

Qy  328 EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 387
         | :| :|| :|: |||:|  |  ||: :| | | |:||||:|| ||| | |||: |:|| |
Db  302 EMITGLDLIAEQIRIAQGEALR--FRQADIQLRGHAIECRINAEDPEYNFRPNPGRITG 358

Qy  388 YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 447
```

```
Query Match          17.1%;    Score 1005.5;   DB 1;   Length 453;
Best Local Similarity 47.1%;   Pred. No. 2.7e-71;
Matches 209;  Conservative 72;  Mismatches 158;  Indels 5;  Gaps 3;

Qy   29 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 88
         | |::||||||||| | |   ::| :|  ||| ::|| ||:|||| |||: ::  | :|
Db    3 FNKILIANRGEIALRLRTCEELGIGTIAVHSTVDRNALHVQLADEAVCIG-EAASSKSY 61

Qy   89 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 148
         |:| ||| || ::| ||:||||||:|||| |||||:|| |||| ||:|  |||| ||
Db   62 LNIPNIIAAALTRNASAIHPGYGFLAENARFAEICADHHLTFIGPSPDSIRAMGDKSTAK 121

Qy  149 TAAKKAGLPVLAESTP-SKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDELRKL 207
         :   ||:| | :|| | ||| |||  |:|||::| ||| ||||||||||:| | |:|||
Db  122 ETMQRVGVPTIPGSDGLLTDVDSAAKVAAEIGYPVMIKATAGGGRGMRLVREPADLEKL 181

Qy  208 ATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVE 267
         ::  | ||||||| | :|:| :|:||:|||::||| || ||||| |||||:|||||||:|
Db  182 FLAAQGEAEAAFGNPGLYLEKFIDRPRHVEFQILADAYGNVVHLGERDCSIQRRHQKLLE 241

Qy  268 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 327
         :|   |||| || | :|| ||| |:||| |||||||||| |||  |||||:|||||| ||
Db  242 EAPSPALSADLRQKMGDAAVKVAQAIGYIGAGTVEFLVDATGNFYFMEMNTRIQVEHPVT 301

Qy  328 EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 387
         | :| :|| :|: |||:|  |  ||: :| | | |:||||:|| ||| | |||: |:|| |
Db  302 EMITGLDLIAEQIRIAQGEALR--FRQADIQLRGHAIECRINAEDPEYNFRPNPGRITG 358

Qy  388 YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 447
         | || ||||:|   ||| |||||||||:| |  |||| || ||||||||| |||:|||
Db  359 YLPGGPGVRVDSHVYTDYEIPPYYDSLIGKLIVWGATREEAIARMQRALRECAITGLPT 418

Qy  448 NIGFLRALLREEDFTSKRIATGFI 471
         ||||:|| |::|| ||| |||||
Db  419 TLSFHQLMLQMPEFLRGELYTNFV 442

RESULT 13
```

```
Db  359 YLPPGGPGVRVDSHVTDYEIPPYYDSLIGKLIVWGATREEAIARMQRALRECAITGLPT 418
Qy  448 NIGFLRALLREEDFTSKRIATGFI 471
Db  419 TLSFHQLMLQMPEFLRGELYTNFV 442
```

RESULT 14
US-08-468-793-8
; Sequence 8, Application US/08468793
; Patent No. 6177267
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert
; APPLICANT: Gornicki, Piotr
; TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468.793
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/422.560
; FILING DATE: 14-APR-1995
; APPLICATION NUMBER: US SN 07/956,700
; FILING DATE: 02-OCT-1992
; CLASSIFICATION: 800
; APPLICATION NUMBER: PCT/US93/09340
; FILING DATE: 30-SEP-1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: ARCD:152/KIT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 453 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-793-8

```
Query Match         17.1%;  Score 1005.5;  DB 3;  Length 453;
Best Local Similarity 47.1%;  Pred. No. 2.7e-71;
Matches  209;  Conservative 72;  Mismatches 158;  Indels 5;  Gaps 3;

Qy   29 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 88
Db    3 FNKILIANRGEIALRLRTCEELGIGTIAVHSTVDRNALHVQLADEAVCIG-EAASSKSY 61
Qy   89 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 148
Db   62 LNIPNIIAALTRNASAIHPGYGFLAENARFAEICADHHLTFIGPSPDSIRAMGDKSTAK 121
Qy  149 TAAKKAGLPVLAESTP-SKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDELRKL 207
```

```
Db  122 ETMQRVGVPTIPGSDGLLTDVDSAAKVAAEIGYPVMIKATAGGGGRGMRLVREPADLEKL 181
Qy  208 ATEASREAEAAFGDGSVYVERAVINPQHIEVQIIGDRTGEVVHLYERDCSLQRRHQKVVE 267
Db  182 FLAAQGEAEAAFGNPGLYLEKFIDRPHVEFQILLADAYGNVHLGERDCSIQRRHQKLLE 241
Qy  268 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHPVT 327
Db  242 EAPSPALSADLRQKMGDAAVKVAQAIGYIGAGTVEFLVDATGNFYFMEMNTRIQVEHPVT 301
Qy  328 EEVTEVDLVKAQMRLAAGATLKELGTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 387
Db  302 EMITGLDLIAEQIRIAQGEALR--FRQADIQLRGHAIECRINAEDPEYNFRPNPGRITG 358
Qy  388 YRSPGGAGVRLDGAAQLGGETAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 447
Db  359 YLPPGGPGVRVDSHVTDYEIPPYYDSLIGKLIVWGATREEAIARMQRALRECAITGLPT 418
Qy  448 NIGFLRALLREEDFTSKRIATGFI 471
Db  419 TLSFHQLMPEFLRGELYINFV 442
```

RESULT 15
US-09-433-043B-121
; Sequence 121, Application US/09433043B
; Patent No. 6399342
; GENERAL INFORMATION:
; APPLICANT: HASELKORN, ROBERT
; APPLICANT: GORNICKI, PIOTR
; TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-CoA CARBOXYLASE
; FILE REFERENCE: ARCD:338US
; CURRENT APPLICATION NUMBER: US/09/433,043B
; CURRENT FILING DATE: 1999-10-25
; PRIOR APPLICATION NUMBER: 08/475,879
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 07/956,700
; PRIOR FILING DATE: 1992-10-02
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 121
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-433-043B-121

```
Query Match         17.1%;  Score 1003.5;  DB 4;  Length 453;
Best Local Similarity 47.1%;  Pred. No. 3.9e-71;
Matches  209;  Conservative 72;  Mismatches 158;  Indels 5;  Gaps 3;

Qy   29 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 88
Db    3 FNKILIANRGEIALRLRTCEELGIGTIAVHSTVDRNALHVQLADEAVCIG-EAASSKSY 61
Qy   89 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 148
Db   62 LNIPNIIAALTRNASAIHPGYGFLAENARFAEICADHHLTFIGPSPDSIRAMGDKSTAK 121
Qy  149 TAAKKAGLPVLAESTP-SKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDELRKL 207
Db  122 ETMQRVGVPTIPGSDGLLTDVDSAAKVAAEIGYPVMIKATAGGGGRGMRLVREPADLEKL 181
Qy  208 ATEASREAEAAFGDGSVYVERAVINPQHIEVQIIGDRTGEVVHLYERDCSLQRRHQKVVE 267
Db  182 FLAAQGEAEAAFGNPGLYLEKFIDRPHVEFQILADAYGNVHLGERDCSIQRRHQKLLE 241
Qy  268 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHPVT 327
Db  242 EAPSPALSADLRQKMGDAAVKVAQAIGYIGAGTVEFLVDATGNFYFMEMNTRIQVEHPVT 301
```

```
Qy  328  EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA  387
              |::|.:||. |:|:|.|.|: - |.|:||| ||| |||.| ||
Db  302  EMITGLDLIAEQIRIAQGEALR--FRQADIQLRGHAIECRINAEDPEYNFRPNPGRITG  358

Qy  388  YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT  447
              | ||| ||||:|   ::|||:::|: .|:-|:||.|||| ::|:|
Db  359  YLPPGGPGVRVDSHVYTDYEIPPYYDSLIGKLIVWGATREEAIARMQRALRECAITGLPT  418

Qy  448  NIGFLRALLREEDFTSKRIATGFI  471
              | .:|. :|  :|.|  |.|:
Db  419  TLSFHQLMLQMPEFLRGELYTNFV  442


Search completed: September 24, 2003, 15:43:53
Job time : 22.1332 secs
```

OM protein - protein search, using sw model

```
          GenCore version 5.1.6
          Copyright (c) 1993 - 2003 Compugen Ltd.

Run on:          September 24, 2003, 15:41:22 ; Search time 36.7701 Seconds
                                              (without alignments)
                                              4760.992 Million cell updates/sec

Title:           US-09-974-973A-2
Perfect score:   5865
Sequence:        1 MTAITLGGLLLKGIIILVST........RVVVPAATKVEGGDLIVVVS 1157

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters:          566894

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_AA:*
            1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
            2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
            3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
            4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
            5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
            6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
            7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
            8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
            9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
           10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
           11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
           12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
           13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
           14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
           15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
           16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
           17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
           18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	5865	100.0	1157	10	US-09-974-973-2	Sequence 2, Appli
2	5865	100.0	1157	10	US-09-974-973-4	Sequence 4, Appli
3	5759	98.2	1140	10	US-09-974-973-19	Sequence 19, Appl
4	5759	98.2	1140	10	US-09-738-626-4265	Sequence 4265, Ap
5	5759	98.2	1140	15	US-10-045-072-2	Sequence 2, Appli
6	2489	42.4	1147	9	US-09-815-242-5468	Sequence 5468, Ap
7	2460.5	42.0	1142	9	US-09-815-242-10806	Sequence 10806, A
8	2300	39.2	1073	9	US-09-815-242-12361	Sequence 12361, A
9	1032.5	17.6	447	9	US-09-767-479-6	Sequence 6, Appli
10	1005.5	17.1	453	9	US-09-767-479-8	Sequence 8, Appli
11	979.5	16.7	471	9	US-09-815-242-5215	Sequence 5215, Ap
12	956	16.3	449	9	US-09-815-242-13885	Sequence 13885, A
13	950	16.2	448	9	US-09-815-242-11160	Sequence 11160, A
14	941	16.0	1171	15	US-10-156-761-14226	Sequence 14226, A
15	940	16.0	449	9	US-09-815-242-10330	Sequence 10330, A
16	936.5	16.0	455	9	US-09-815-242-13617	Sequence 13617, A
17	935.5	16.0	455	9	US-09-815-242-13364	Sequence 13364, A
18	921	15.7	449	9	US-09-815-242-12063	Sequence 12063, A
19	919	15.7	458	9	US-09-815-242-11321	Sequence 11321, A
20	912	15.5	590	15	US-10-156-761-10874	Sequence 10874, A
21	909	15.5	443	9	US-09-815-242-4963	Sequence 4963, Ap
22	909	15.5	456	9	US-09-815-242-10924	Sequence 10924, Ap
23	906	15.4	590	12	US-10-045-612A-25	Sequence 25, Appl
24	902	15.4	455	9	US-09-815-242-11558	Sequence 11558, A
25	897	15.3	725	15	US-10-160-501-17	Sequence 17, Appli
26	896	15.3	725	15	US-10-224-539A-2	Sequence 2, Appli
27	896	15.3	725	15	US-10-224-539A-9	Sequence 9, Appli
28	893	15.2	590	12	US-10-045-612A-26	Sequence 26, Appl
29	890	15.2	700	15	US-10-156-761-12811	Sequence 12811, A
30	880.5	15.0	591	10	US-09-738-626-6940	Sequence 6940, Ap
31	856.5	14.6	446	9	US-09-815-242-5418	Sequence 5418, Ap
32	856.5	14.6	453	9	US-09-815-242-12562	Sequence 12562, A
33	837	14.3	448	9	US-09-815-242-5806	Sequence 5806, Ap
34	834	14.2	451	9	US-09-815-242-13127	Sequence 13127, A
35	829.5	14.1	616	15	US-10-156-761-11400	Sequence 11400, A
36	653	11.1	464	15	US-10-169-048-28	Sequence 28, Appl
37	633	10.8	358	9	US-09-815-242-12939	Sequence 12939, A
38	583.5	9.9	2257	9	US-09-767-479-10	Sequence 10, Appl
39	533.5	9.1	2257	12	US-09-839-477-8	Sequence 8, Appli
40	437.5	7.5	483	15	US-10-083-357-1328	Sequence 1328, Ap
41	423.5	7.2	262	15	US-10-224-539A-5	Sequence 5, Appli
42	377.5	6.4	158	9	US-09-815-242-5031	Sequence 5031, Ap
43	300.5	5.1	163	10	US-09-895-913A-324	Sequence 324, Appl
44	252	4.3	1073	16	US-09-210-115-20	Sequence 20, Appl
45	251	4.3	124	9	US-09-205-658-240	Sequence 240, App

ALIGNMENTS

```
RESULT 1
US-09-974-973-2
; Sequence 2, Application US/09974973
; Patent No. US2002017202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacter
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974,973
; CURRENT FILING DATE: 2001-10-21
; PRIOR APPLICATION NUMBER: US 60/239,913
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1157
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-974-973-2

Query Match            100.0%;  Score 5865;  DB 10;  Length 1157;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1157;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy   1 MTAITLGGLLLKGIIILVSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYP 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MTAITLGGLLLKGIIILVSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYP 60

Qy  61 REDRGSFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLA 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 REDRGSFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLA 120

Qy 121 RECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTY 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 RECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTY 180

Qy 181 PIFVKAVAGGGGRGMRFVSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQI 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
;
;   TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacter
;   FILE REFERENCE: 1533.123000I/MAC/RGM
;   CURRENT APPLICATION NUMBER: US/09/974,973
;   CURRENT FILING DATE: 2001-10-21
;   PRIOR APPLICATION NUMBER: US 60/239,913
;   PRIOR FILING DATE: 2000-10-13
;   NUMBER OF SEQ ID NOS: 19
;   SOFTWARE: PatentIn version 3.0
;   SEQ ID NO 4
;   LENGTH: 1157
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-974-973-4

Query Match             100.0%;  Score 5865;  DB 10;  Length 1157;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches  1157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     1  MTAITLGGLLKGIITILVSTHTSSTLPAFKKILVANRGEIAVRFRAALETGAATVAIYP   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MTAITLGGLLKGIITILVSTHTSSTLPAFKKILVANRGEIAVRFRAALETGAATVAIYP   60

Qy    61  REDRGSFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLA  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  REDRGSFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLA  120

Qy   121  RECAENGITFIGPTPEVLLDTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTY  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  RECAENGITFIGPTPEVLLDTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTY  180

Qy   181  PIFVKAVAGGGRGMRFVSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQI   240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  PIFVKAVAGGGRGMRFVSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQI   240

Qy   241  LGDRTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGT  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  LGDRTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGT  300

Qy   301  VEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKT  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  VEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKT  360

Qy   361  HGAALQCRITTEDPNNGFRPDTGITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMT   420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  HGAALQCRITTEDPNNGFRPDTGITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMT   420

Qy   421  CRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQA  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  CRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQA  480

Qy   481  PPADDEQGRILDYLADVTVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAA  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  PPADDEQGRILDYLADVTVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAA  540

Qy   541  FARDLREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATY  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  FARDLREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATY  600

Qy   601  DVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVD  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  DVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVD  660

Qy   661  IFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIV  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  IFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIV  720

Qy   721  KSGAHILAIKDMAGLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYWEAVRGLYL  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  KSGAHILAIKDMAGLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYWEAVRGLYL  780

Qy   781  AVDGASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGT  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  AVDGASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGT  840

Qy   841  PGPTGRYYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVV  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  PGPTGRYYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVV  900

Qy   901  GDLALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRIRALEGRSEG  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  GDLALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRIRALEGRSEG  960

Qy   961  KAPLTEVPEEEQAHLDADDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  KAPLTEVPEEEQAHLDADDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYG 1020

Qy  1021  LVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTAT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  LVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTAT 1080

Qy  1081  AEKADSSNKGHVAAPFAGVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVV 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  AEKADSSNKGHVAAPFAGVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVV 1140

Qy  1141  VPAATKVEGGDLIVVS                                             1157
          |||||||||||||||
Db  1141  VPAATKVEGGDLIVVS                                             1157
```

```
Qy   378  FRPDTGTTPAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  437
          ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  FRPDTGTTTPAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  420

Qy   438  AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADV  497
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   421  AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV  480

Qy   498  TVNKPHGVRPKDVAAPIDKLPLNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVIDTT  557
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  TVNKPHGVRPKDVAAPIDKLPLNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVIDTT  540

Qy   558  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGATYDVAMRFLFEDPWDRLDE  617
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGATYDVAMRFLFEDPWDRLDE  600

Qy   618  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  677
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  660

Qy   678  IDAVLETNTAVAEVAMAYSGDLLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR  737
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  IDAVLETNTAVAEVAMAYSGDLLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR  720

Qy   738  PAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  797
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  PAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  780

Qy   798  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQL  857
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQL  840

Qy   858  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF  917
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF  900

Qy   918  AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA  977
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA  960

Qy   978  DDSKERRNSLNRLLPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP  1037
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  DDSKERRNSLNRLLPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP  1020

Qy  1038  LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA  1097
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA  1080

Qy  1098  GVVTVTVAEGDEVKAGDAVAIIEAMKMEATTTASVDGKIERVVPAATKVEGGDLIVVVS  1157
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db  1081  GVVTVTVAEGDEVKAGDAVAIIEAMKMEATTTASVDGKIDRVVPAATKVEGGDLIVVVS  1140
```

RESULT 4
US-09-738-626-4265
; Sequence 4265, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;   APPLICANT: NAKAGAWA, SATOSHI
;   APPLICANT: MIZOGUCHI, HIROSHI
;   APPLICANT: ANDO, SEIKO
;   APPLICANT: HAYASHI, MIKIRO
;   APPLICANT: OCHIAI, KEIKO
;   APPLICANT: YOKOI, HARUHIKO
;   APPLICANT: TATEISHI, NAOKO
;   APPLICANT: SENOH, AKIHIRO
;   APPLICANT: IKEDA, MASATO
;   APPLICANT: OZAKI, AKIO
;   TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;   FILE REFERENCE: 249-125
;   CURRENT APPLICATION NUMBER: US/09/738,626
;   CURRENT FILING DATE: 2000-12-18

```
Qy   841  PGPTGRVYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVV  900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  PGPTGRVYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVV  900

Qy   901  GDLALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEG  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  GDLALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEG  960

Qy   961  KAPLTEVPEEEQAHLDADDSKERRNSLNRLLPKPTEEFLEHRRRFGNTSALDDREFFYG  1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  KAPLTEVPEEEQAHLDADDSKERRNSLNRLLPKPTEEFLEHRRRFGNTSALDDREFFYG  1020

Qy  1021  LVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTAT  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  LVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTAT  1080

Qy  1081  AEKADSSNKGHVAAPFAGVVTVTVAEGDEVKAGDAVAIIEAMKMEATTTASVDGKIERVV  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  AEKADSSNKGHVAAPFAGVVTVTVAEGDEVKAGDAVAIIEAMKMEATTTASVDGKIERVV  1140

Qy  1141  VPAATKVEGGDLIVVVS  1157
          |||||||||||||||||
Db  1141  VPAATKVEGGDLIVVVS  1157
```

RESULT 3
US-09-974-973-19
; Sequence 19, Application US/09974973
; Patent No. US2002017202A1
; GENERAL INFORMATION:
;   APPLICANT: Hanke, Paul D.
;   TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacterium
;   FILE REFERENCE: 1533.1230001/MAC/RGM
;   CURRENT APPLICATION NUMBER: US/09/974,973
;   CURRENT FILING DATE: 2001-10-21
;   PRIOR APPLICATION NUMBER: US 60/239,913
;   PRIOR FILING DATE: 2000-10-13
;   NUMBER OF SEQ ID NOS: 19
;   SOFTWARE: PatentIn version 3.0
;   SEQ ID NO 19
;   LENGTH: 1140
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-974-973-19

```
Query Match        98.2%;  Score 5759;  DB 10;  Length 1140;
Best Local Similarity  99.48%;  Pred. No. 0;
Matches 1133;  Conservative 5;  Mismatches 2;  Indels 0;  Gaps 0;

Qy    18  VSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR  77
          :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR  60

Qy    78  IGTEGSPVKAYLDIDEIIGAAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  137
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  IGTPGSPVKAYLDIDEIIGAAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  120

Qy   138  LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRF  197
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   121  LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF  180

Qy   198  VSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCS  257
          |:||||||||||||||||||||||| ||||||||||||||||||||| ||||||||||||
Db   181  VASPDELRKLATEASREAEAAFGDGAVVYERAVINPQHIEVQILGDHTGEVVHLYERDCS  240

Qy   258  LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN  317
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN  300

Qy   318  PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG  377
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   301  PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLIQDKIKTHGAALQCRITTEDPNNG  360
```

```
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4265
;   LENGTH: 1140
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-4265

Query Match      98.2%;  Score 5759;  DB 10;  Length 1140;
Best Local Similarity  99.4%;  Pred. No. 0;
Matches 1133;  Conservative 5;  Mismatches 5; Indels 2;  Gaps  0;

Qy   18  VSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR   77
         :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR   60

Qy   78  IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  137
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  120

Qy  138  LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRF  197
         ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db  121  LDLTGDKSRAVTAAKKAGLPVLAESTPSKNDEIVKSAEGQTYPIFVKAVAGGGGRGMRF  180

Qy  198  VSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCS  257
         |:|||||||||||||||||||||||:|||||||||||||||||||:||||||||||||||
Db  181  VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS  240

Qy  258  LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN  317
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN  300

Qy  318  PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG  377
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG  360

Qy  378  FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  437
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  420

Qy  438  AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADV  497
         |||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db  421  AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV  480

Qy  498  TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  557
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  540

Qy  558  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  617
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  600

Qy  618  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  677
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  660

Qy  678  IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR  737
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR  720

Qy  738  PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  797
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  780

Qy  798  LSAIAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPESGTPGTGRVYRHEIPGGQL  857
         |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPESGTPGTGRVYRHEIPGGQL  840

Qy  858  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF  917
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF  900

Qy  918  AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRIRALEGRSEGKAPLTEVPEEQAHLDA  977
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRIRALEGRSEGKAPLTEVPEEQAHLDA  960

Qy  978  DDSKERRSNLRLLFPKPTEEFLEHRRRFGNTSALDDRFFYGLVEGRETLIRLPDVRTP 1037
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  DDSKERRSNLRLLFPKPTEEFLEHRRRFGNTSALDDRFFYGLVEGRETLIRLPDVRTP 1020

Qy 1038  LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1097
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080

Qy 1098  GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVVPAATKVEGGDLIVVVS 1157
         |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db 1081  GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS 1140


RESULT 5
US-10-045-072-2
; Sequence 2, Application US/10045072
; Publication No. US2003002730SA1
; GENERAL INFORMATION:
;   APPLICANT: Sinskey, Anthony J.
;   APPLICANT: Lessard, Philip A.
;   APPLICANT: Willis, Laura B.
;   TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
;   FILE REFERENCE: 1533.0790002
;   CURRENT APPLICATION NUMBER: US/10/045,072
;   CURRENT FILING DATE: 2002-01-15
;   PRIOR APPLICATION NUMBER: US 09/677,575
;   PRIOR FILING DATE: 2000-10-03
;   PRIOR APPLICATION NUMBER: US 09/220,081
;   PRIOR FILING DATE: 1998-12-23
;   NUMBER OF SEQ ID NOS: 39
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1140
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-10-045-072-2

Query Match      98.2%;  Score 5759;  DB 15;  Length 1140;
Best Local Similarity  99.4%;  Pred. No. 0;
Matches 1133;  Conservative 5;  Mismatches 2; Indels 0;  Gaps  0;

Qy   18  VSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR   77
         :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR   60

Qy   78  IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  137
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  120

Qy  138  LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRF  197
         ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db  121  LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF  180

Qy  198  VSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCS  257
         |:|||||||||||||||||||||||:|||||||||||||||||||:||||||||||||||
Db  181  VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS  240

Qy  258  LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN  317
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN  300

Qy  318  PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG  377
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG  360
```

```
;   PRIOR FILING DATE: 2000-03-21
;   PRIOR APPLICATION NUMBER: 60/206,848
;   PRIOR FILING DATE: 2000-05-23
;   PRIOR APPLICATION NUMBER: 60/207,727
;   PRIOR FILING DATE: 2000-05-26
;   PRIOR APPLICATION NUMBER: 60/242,578
;   PRIOR FILING DATE: 2000-10-23
;   PRIOR APPLICATION NUMBER: 60/253,625
;   PRIOR FILING DATE: 2000-11-27
;   PRIOR APPLICATION NUMBER: 60/257,931
;   PRIOR FILING DATE: 2000-12-22
;   PRIOR APPLICATION NUMBER: 60/269,308
;   PRIOR FILING DATE: 2001-02-16
;   NUMBER OF SEQ ID NOS: 14110
;   SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 5468
;   LENGTH: 1147
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-09-815-242-5468

  Query Match      42.4%;  Score 2489;  DB 9;  Length 1147;
  Best Local Similarity  45.2%;  Pred. No. 3.1e-190;
  Matches  527;  Conservative 193;  Mismatches 383;  Indels 62;  Gaps 14;

Qy    30  KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL   89
Db     4  KKLLVANRGEIAIRIFRAAAELDISTVAIYSNEDKSSLHRYKADESYLVGSDLGPAESVL   63

Qy    90  DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT  149
Db    64  NIERIIDVAKQANVDAIHPGYGFLSENEQFARRCAEEGIKFIGPHLEHLDMFGDKVKART  123

Qy   150  AAKKAGLPVL-AESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDELRKIA  208
Db   124  TAIKADLPVIPGTDGPIKSYELAKEFAEEAGFPLMIKATSGGGGKGMRIVREESELEDAF  183

Qy   209  TEASREEAAFGDGSVYVERAVINPQHIEVQLLGDRTGEVVHLYERDCSLQRRHQKVVEI  268
Db   184  HRAKSEAKSFGNSEVYIERYIDNPKHIEVQVGDEHGNIVHLFERDCSVQRRHQKVVEV  243

Qy   269  APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLV--DEKGNHVFIEMNPRIQVERTV  326
Db   244  APSVGLSPTLRQRICDAAIQLMENIKYNAGIVEFLVSGDE--FFFIEVNPRVQVEHTI  300

Qy   327  TEEVTEVDLVKAQMRLAAGATL-KELGLTQDK-IKTHGAALQCRITTEDPNNGFRPDTG  383
Db   301  TEMVTGIDIVKTQILVAAGADLFGEEINMPQQKDITTLGYAIQCRITTEDPLNDFMPDTG  360

Qy   384  TITTAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTV  442
Db   361  TIIAYRSSGGFGVRLDAGDGFQGAEISPYYDSILVKLSTHAISFKQAEEKMVRSLREMRI  420

Qy   443  SGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKP  502
Db   421  RGVKTNIPFLINVMKNKKFTSGDYTTKFIEETPELFDIQPSLDRGTKILEYIGNVTIN--  478

Qy   503  HGVRPKDVAAPIDKLPNIKDLPLP------------RGSRDRLKQLGPAAFARD  544
Db   479  ---------GFPNVEKRPKPDYELASIPTVSSSKIASFSGTKQLLDEVGPKGVAEW  525

Qy   545  LREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAM  604
Db   526  VKKQDDVLLTDTTFRDAHQSLLATRVRTKDMINIASKTADVFKDGFSLEMWGGATFDVAY  585

Qy   605  RPLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRI  664
Db   586  NFLKENPWERLERLRKAIPNVLFQMLLRASNAVGYKNYPDNVIHKFVQESAKAGIDVFRI  645

Qy   665  FDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDP-NEKLYTLDYLKMAEEIVKSG  723
Db   646  FDSLNWVDQMKVANEAVQEAG-KISEGTICYTGDILNPERSNIYTLEYYVKLAKELEREG  704
```

```
Qy   378  FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  437
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  420

Qy   438  AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADV  497
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV  480

Qy   498  TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  557
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  540

Qy   558  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  617
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  600

Qy   618  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  677
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  660

Qy   678  IDAVLETNTAVAEVAMAYSGDLSDPNEKLLYTLDYLKMAEEIVKSGAHILAIKDMAGLLR  737
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  IDAVLETNTAVAEVAMAYSGDLSDPNEKLLYTLDYLKMAEEIVKSGAHILAIKDMAGLLR  720

Qy   738  PAAVKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  797
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  PAAVKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  780

Qy   798  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQL  857
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQL  840

Qy   858  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF  917
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF  900

Qy   918  AADPQKYDIPDSVIAFLRGELGNPEGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA  977
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  AADPQKYDIPDSVIAFLRGELGNPEGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA  960

Qy   978  DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP  1037
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP  1020

Qy  1038  LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA  1097
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA  1080

Qy  1098  GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVPAATKVEGGDLIVVVS  1157
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS  1140
```

```
RESULT 6
US-09-815-242-5468
;   Sequence 5468, Application US/09815242
;   Patent No. US2002006156A1
;   GENERAL INFORMATION:
;   APPLICANT: Haselbeck, Robert
;   APPLICANT: Ohlsen, Karl L.
;   APPLICANT: Zyskind, Judith W.
;   APPLICANT: Wall, Daniel
;   APPLICANT: Trawick, John D.
;   APPLICANT: Carr, Grant J.
;   APPLICANT: Yamamoto, Robert T.
;   APPLICANT: Xu, H. Howard
;   TITLE OF INVENTION: Identification of Essential Genes in
;   TITLE OF INVENTION:  Prokaryotes
;   FILE REFERENCE: ELITRA.011A
;   CURRENT APPLICATION NUMBER: US/09/815,242
;   CURRENT FILING DATE: 2001-03-21
;   PRIOR APPLICATION NUMBER: 60/191,078
```

```
Query Match      42.0%;   Score 2460.5;   DB 9;   Length 1142;
Best Local Similarity  45.5%;   Pred. No. 5.9e-188;
Matches  529;   Conservative 187;   Mismatches 387;   Indels 59;   Gaps 17;

Qy   30 KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL  89
         ||:|||||||||:| :|| ||| ||||:| | ||   | :| ||  :|| |    |||||
Db    2 KKVLVANRGEIAIRIFRACTELDIRTVAIYAAEDEYSVHRFKADEAYLVGKEKKPIEAYL  61

Qy   90 DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGTPEVLDLIGDKSRAVT  149
         || :|| |||| ||||:|||||||||  : |  |   ||||:| | ||   ||||  | |
Db   62 DIENIIQIAKKSGADAIHPGYGFLSENLRFAERCEEEGIIFVGPKTHHLDIFGDKIKAKE  121

Qy  150 AAKKAGLPVLAEST-PSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKLA  208
         || :||: :  |:  |    | ||| |      :||||  ||||||||| :    :|||||
Db  122 AAVAAGIASIPGSDGPVATVEEVVAFGETHGFPIMIKAALGGGGRGMRVAHDAKEAREGY  181

Qy  209 TEASREAAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEI  268
         ::| ||| ||| ||:||| ::|||:||||||:||  | |:||:|||||:|||||||||:
Db  182 ERAKSEAKAAFGSDEVYVEKYISNPKHIEVQIIGDHGNVLHLFERDCSVQRRHQKVVEV  241

Qy  269 APAQHLDPELRDRICADAVKFCRSIGYQGAGVEFLVDEKGNHVFIEMNPRIQVEHTVTE  328
         || :  ::   :|||:||| :   :|| ||||||||  ::  :|||:|||:|||||:||
Db  242 APCVSMNEEQRAAICSAAVQLMAHVGVNAGTVEFLV-EGDQFYFIEVNPRVQVEHTITE  300

Qy  329 EVTEVDLVKAQMRLAAGATL-KELGL-TQDKIKTHGAALQCRITTEDPNNGFRPDTGTIT  386
          :| :| : ||:| |  | | |: |   ::|:  ||||:||||||||| |||| ||| | |
Db  301 MITDIDIVISQLQIAQGLDLHKDMHLPKQNELTLKGAAIQCRITTEDPLNQFMPDTGKID  360

Qy  387 AYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGV  445
          |||||  ||||| |||  ||  |  ||| ||||   :| || | | :||| |||  |||
Db  361 TYRSPGFGVRLDVGNAYSGYAVTPYFDSLLVKVCTHGFSFEQAISKMQRCLKEFRIRGV  420

Qy  446 ATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKPHGV  505
          :||| || :   |:|||  :||| |||||||     |||:| | | | ||| |||| | |
Db  421 KTNIPFLQNVSYPAFQSGEAKTTFIDNTPELFEFPMRDRGNKTMKYIGEVTVNGFPGi  480

Qy  506 R---------PKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAV  553
                  |:  || |||:|| | |   | |   ||| |:|   || |||
Db  481 ERTEKKYFEAPRVPTDIEVP-EKVITAKNI---------LDAQGATAVIDWVKNQESVLM  530

Qy  554 TDTTFRDAHQSLLATRVRSFALKPAAEAVAKLT---PELLSVEAWGGATYDVAMRFLFE  609
         ||||||||||||||||||:   ||      |:    ||| | | ||||| |||  ||::
Db  531 TDTTFRDAHQSLLATRVRTQDFK----ALAGLTDVALPELFSSEMWGGATFDVAYRFLTE  586

Qy  610 DPWDRLDELRERAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALN  669
         |||:||  :|: :||| ::|||:||:| |||  |||.| :| :||:| |||:|||||| ||
Db  587 DPWQRLRKIRQLMPNILLQMLFRGSNAVGYQNYPDNVIEEFIKESARQGVDVFRIFDSLN  646

Qy  670 DVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAI  729
          :  | |    :||:|  |  ||  | || :||  |||   ||  |||:  | |||:|||
Db  647 WIPQMEKSIQVVRDTG-KIAEAICYTGDINDPARAKYNVQYLDMAKELENLGAQIIAI  705

Qy  730 KDMAGLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPL  789
         |||||||:| | ::| :||: |:||:|:|||||:| :|: || ||  |||| |  ||
Db  706 KDMAGLLKPQAAYRLISELKAATDLPIHLHTHDTSGNGIITYSAATKAGVDIVDVAMSAM  765

Qy  790 SGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYR  849
         || :||| ::::::|  :  |: :  |:   :  | |  |||:|: || | |  | ||
Db  766 SGATSQPSMNSLYYALVNGERTPTINIDNAQKINHYWEDVRMYYQPFENGLNAPQTEVYM  825

Qy  850 HEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVG  909
         ||:||||:||| :: | ||| :::|: ::||  :|:|:|| ::|||||||||||:||:|
Db  826 HEMPGGQYSNLQQQAKAVGLGHRWDEIKKMYHTVNLMFGDIVKVTPSSKVVGDMALFMVQ  885

Qy  910 AGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPE  969
          |:   ||    ::| |||||: | :|:|| | ||:|:|: :|  |      ||:||
Db  886 NHLTEQDVYARGEELSFPESVVTFQGDLGQPVGGFPKELQRIILKGR---PAFTERPG  941

Qy  970 EEQAHLDADDSKER---------RNSLNRLLFPKPTEEFLEHRRR--FGNTSALDDR  1015
         :   ::: :: :|       : ::   |  :   ::|  | : |   | :   ||
Db  942 DLAAPVDFAKVQEELAEKIGYQPKLEEVLSYLMYP--QVFLEYRQKYETFGDITLLDTP  998

Qy 1016 EFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVE  1075
```

```
Qy  724 AHILAIKDMAGLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVD  783
         |||||||||||||:| : |||| ||||:|||:|:|||||:||: | |  ||  ||||||
Db  705 FHILAIKDMAGLLKPKAAYELIGELKAAVDLPIHLHTHDTSGNGLLTYKQAIDAGVDIID  764

Qy  784 GASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGP  843
           |:  ||||:||| ::::| | :  |:|:|||:     | | | |||:| || |||
Db  765 TAVASMSGLTSQPSANSLYYALNGFPRHLRTDIEGMESLLSHYWSTVRTYYSDFESDIKSP  824

Qy  844 TGRVYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDL  903
         ||:|:||:||||| |||:::   |||| ||: :||:|:  :| :|| :||||||||||||
Db  825 NTEIYQHEMPGGQYSNLSQQAKSLGLGERFDEVKDMYRRVNFLFGDIVKVTPSSKVVGDM  884

Qy  904 ALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAP  963
         || :| :::||  |  |   | : |||::|| |:|| | || |: |: |        |
Db  885 ALYMVQNDLDEQSVITDGYKLDFPPESVVSFFKGEIGQPVNGFNKDLQAVILKGQE---A  940

Qy  964 LTEVPEEEQAHLDADDSKE--------RRNSLNRLLFPKPTEEFLEHRRRFGNTSAL  1012
         ||   |:|   :::  ::         :|:   :|   |::: :::   |  ::| ||
Db  941 LTARPGEYLEPVDFEKVRELLEEEQQGPVTEQDIISYVLYPKVVEQYIQTRNQYGNLSLL  1000

Qy 1013 DDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDR  1072
         | :  :||: ||| |  |:|: :|:|::|:::||||| :|:|  :  | |||  |:::
Db 1001 DTPTFFGMRNGETVEIEI-DKGKRLIIKLETISEPDENGNRTIYYAMNGQARRIYIKDE  1059

Qy 1073 SVESVTATAEKADSSNKGHVAAPPAGVVT-VTVAEGDEVKAGDAVAIITAMKMEATITAS  1131
          | ||| || :|:||| ||: | :  |:| :||| ||||||||||||||||||| || :
Db 1060 NVHTNANVKPKADKSNPSHIGAQMPGSVTEVKVSVGETVEVKAGDAVAIITAMKMETTIQAP  1119

Qy 1132 VDGKIERVVVPAAATKVEGGDLIVVV  1156
         .|| |::| |  | |:|:|||:|::
Db 1120 FDGVIKQVTVNNGDTIATGDLLIFI  1144
```

RESULT 7
US-09-815-242-10806
; Sequence 10806, Application US/09815242
; Patent No. US2002006l569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10806
; LENGTH: 1142
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10806

```
Db   999 TFFNGIRQGETLEVQIERGKI-LIIRLDEIGEPDIDGNRVLFFNLNGQRREVLVKDASIK 1057
Qy  1076 SVTATAEKADSSNKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDG 1134
Db  1058 SAVQVKQKAEPTNKEQIGAIMSGSVLQVLVKRGDKVEKGQPLLITEAMKMETTIEARFAG 1117
Qy  1135 KIERVVVPAATKVEGGDLIVVV 1156
Db  1118 TVDHIYVEEGEAISSGDLLLEV 1139

RESULT 8
US-09-815-242-12361
; Sequence 12361, Application US/09815242
; Patent No. US2002006156941
GENERAL INFORMATION:
  APPLICANT: Haselbeck, Robert
  APPLICANT: Ohlsen, Kari L.
  APPLICANT: Zyskind, Judith W.
  APPLICANT: Wall, Daniel
  APPLICANT: Trawick, John D.
  APPLICANT: Carr, Grant J.
  APPLICANT: Yamamoto, Robert T.
  APPLICANT: Xu, H. Howard
  TITLE OF INVENTION: Identification of Essential Genes in
  TITLE OF INVENTION: Prokaryotes
  FILE REFERENCE: ELITRA.011A
  CURRENT APPLICATION NUMBER: US/09/815,242
  CURRENT FILING DATE: 2001-03-21
  PRIOR APPLICATION NUMBER: 60/191,078
  PRIOR FILING DATE: 2000-03-21
  PRIOR APPLICATION NUMBER: 60/206,848
  PRIOR FILING DATE: 2000-05-23
  PRIOR APPLICATION NUMBER: 60/207,727
  PRIOR FILING DATE: 2000-05-26
  PRIOR APPLICATION NUMBER: 60/242,578
  PRIOR FILING DATE: 2000-10-23
  PRIOR APPLICATION NUMBER: 60/253,625
  PRIOR FILING DATE: 2000-11-27
  PRIOR APPLICATION NUMBER: 60/257,931
  PRIOR FILING DATE: 2000-12-22
  PRIOR APPLICATION NUMBER: 60/269,308
  PRIOR FILING DATE: 2001-02-16
  NUMBER OF SEQ ID NOS: 14110
  SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 12361
  LENGTH: 1073
  TYPE: PRT
  ORGANISM: Staphylococcus aureus
US-09-815-242-12361

Query Match   39.2%;  Score 2300; DB 9;  Length 1073;
Best Local Similarity  44.8%;  Pred. No. 4e-175;
Matches 489;  Conservative 180;  Mismatches 360;  Indels 62;  Gaps 14;

Qy  104 DAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVL-AES 162
Db    2 DATHPGYGFLSENEQFARRCAEEGIKFIGPHLEHLDMFGDKVKARTTAIKADLPVIPGTD  61
Qy  163 TPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKLATEASREABAAFGDG 222
Db   62 GPIKSYELAKEFAEEAGFPLMIKATSGGGGKGMRIVREESELEDAFHRAKSEAEKSFGNS 121
Qy  223 SVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRI 282
Db  122 EVYIERYIDNPKHIEVQVIGDEHGNIVHLFERDCSVQRRHQKVVEVAPSVGLSPTLRQRI 181
Qy  283 CADAVKFCRSIGYQGAGTVEFLV-DEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQM 340
Db  182 CDAAIQLMENIKYYNAGTVEFLVSGDE---FFFIEVNPRVQVEHTITEMVTGIDIVKTQI 238
Qy  341 RLAAGATL--KELGLTQDK-IKTHGAALQCRITTEDPNNGFRPDTGTITAYRSPGGAGVR 397
Db  239 LVAAGADLFGEEINPQQKDITTLGYALQCRITTEDPLNDFMPDTGTTIAYRSSGFGVR   298
Qy  398 LD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIGFLRALL 456
Db  299 LDAGDFQGAEISPYYDSLLVKLLSTHAISFKQAEEKMVRSLREMRIRGVKTNIPFLINVM 358
Qy  457 REEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILLDYLADVTVNKPHGVRPKDVAAPIDK 516
Db  359 KNKKFTSGDYTKFIEETPELFDIQPSLDRGTKTLEYIGNVTIN-------------G 403
Qy  517 LPNIKDLPLP---------------RGSRDRLKQLGPAAFARDLREQDALAVTDTTF 558
Db  404 FPNVEKRPKPDYELASIPTVSSSKIASFSGTKQLLDEVGPKGVAEWVKKQDDVLLTDTTF 463
Qy  559 RDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRLFEDPWDRLDEL  618
Db  464 RDAHQSLLATRVRTKDMINIASKTADVFKDGFSLEMWGGATFDVAYNFLKENPWERLERI 523
Qy  619 REAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAI 678
Db  524 RKAIPNVLFQMLLRASNAVGYKNYPDNVIHKFVQESAKAGIDVFRIFDSLNWVDQMKVAN 583
Qy  679 DAVLETNTAVAEVAMAYSGDLSDP-NEKLYTLDYLKMAEEIVKSGAHILAIKDMAGLLR  737
Db  584 EAVQEAG-KISEGTICYTGDILNPERSNIYTLEYYVKLAKELEREGFHILAIKDMAGLLK 642
Qy  738 PAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  797
Db  643 PKAAYELIGELKAAVDLPIHLHTHDTSGNGLLFYKQAIDAGVDIIDTAVASMSGLTSQPS 702
Qy  798 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 857
Db  703 ANSLYYALNGPFRHLRTDIEGMESLSHWSTVRTYYSDFESDIKSPNTEIYQHEMPGGQY  762
Qy  858 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVYGDLALHLVGAGVDPADF 917
Db  763 SNLSQQAKSLGLGERFDEVKDMYRRVNFLFGDIVKVTPSSKVVGDMALYMVQNDLDEQSV 822
Qy  918 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDA  977
Db  823 ITDGYKLDFPESVVSFFKGEIGQPVNGFNKDLQAVILKGQE---ALTARPGEYLEPVDF  878
Qy  978 DDSKE------------RRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRE 1026
Db  879 EKVRELLEEEQQGPVTEQDIISYLVLYPKVYEQYIQTRNQYGNLSLLDTPTFFGMRNGET 938
Qy 1027 TLIRLPDVRTPLLVRLLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADS 1086
Db  939 VEIEI-DKGKRLIIKLETISEPDENGNRTIYAMNGQARRTYIKDENVHTNANVKPKADK  997
Qy 1087 SNKGHVAAPFAGVVT-VTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVVPAAT 1145
Db  998 SNPSHIGAQMPGSVTEVKVSVGETVKANQPLLITEAMKMETTIQAPFDGVIKQVTVNNGD 1057
Qy 1146 KVEGGDLIVVV 1156
Db 1058 TIATGDLLIEI 1068

RESULT 9
US-09-767-479-6
; Sequence 6, Application US/09767479
; Patent No. US20010036544A1
GENERAL INFORMATION:
  APPLICANT: Haselkorn, Robert
  APPLICANT: Gornicki, Piotr
  TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
                      METHODS OF USE
  NUMBER OF SEQUENCES: 29
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Arnold, White & Durkee
    STREET: P.O. Box 4433
```

```
          CITY: Houston
          STATE: Texas
          COUNTRY: United States of America
          ZIP: 77210
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
        SOFTWARE: PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/09/767,479
        FILING DATE: 22-Jan-2001
        CLASSIFICATION: Unknown
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: 08/468,793
        FILING DATE: <Unknown>
        APPLICATION NUMBER: US SN 07/956,700
        FILING DATE: 02-OCT-1992
        APPLICATION NUMBER: PCT/US93/09340
        FILING DATE: 30-SEP-1993
    ATTORNEY/AGENT INFORMATION:
        NAME: Kitchell, Barbara S.
        REGISTRATION NUMBER: 33,928
        REFERENCE/DOCKET NUMBER: ARCD:152/KIT
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: (512) 418-3000
        TELEFAX: (713) 789-2679
        TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 6:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 447 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-767-479-6

Query Match             17.6%;  Score 1032.5;  DB 9;  Length 447;
Best Local Similarity   48.2%;  Pred. No. 4.1e-74;
Matches  216;  Conservative  71;  Mismatches  148;  Indels  13;  Gaps  4;

Qy   29  FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY  88
            ||||||||||||  |  |  | :|  || :|
Db    3  FDKILIANRGEIALRILRACEEMGIATIAVHSTVDRNALHVQLADEAVCIG-EPASAKSY  61

Qy   89  LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLIGDKSRAV  148
            |:|||||||||||||||||||||||  |  :|||| | | |  | | |  |||
Db   62  LNIPNIIAAALTRNASAIHPGYGFLSENAKFAEICADHHIAFIGPTPEAIRLMGDKSTAK  121

Qy  149  TAAKKAGLPVLAES-----TPSKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDE  203
            ||||  | :|:|        :|||| |:||:|| ||||||:|||||||
Db  122  ETMQKAGVPTVPGSEGLVETEQEGLE---LAKDIGYPVMIKATAGGGGRGMRLVRSPDE  177

Qy  204  LRKLATEASREAEAAFGDGSVYYERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQ  263
            |:| | |  | :|| ||:|||:|||||:|||||:||:|||
Db  178  FVKLFLAAQGEAGAAFGNAGVYIEKFIERPRHIEFQILADNYGNVIHLGERDCSIQRRNQ  237

Qy  264  KVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVE  323
            |::| | |  |  ||:||| | :|| ||||:|||
Db  238  KLLEEAPSPALDSDLREKMGQAAVKAAQFINYTGAGTIEFLLDRSGQFYFMEMNTRIQVE  297

Qy  324  HTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTG  383
            | ||| ||  |||| :| ||||   ||| |||| |:| ||:| |||||||:|| |||  |
Db  298  HPVTEMVTGVDLLVEQIRIAQGERLR---LTQDQVVLRGHAIECRINAEDPDHDFRPAPG  354

Qy  384  TITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVS  443
            | :  :||| |||:| | |  |  ||:| ||:|||| |||| ||||||:||||| |  :|
Db  355  RISGYLPPGGPGVRIDSHVYTDYQIPPYYDSLLGKLIVWGPDRATAINRMKRALRECAIT  414

Qy  444  GVATNIGFLRALLREEDFTSKRIATGFI  471
            | : |||| :| : ::|| |: :::||:|
Db  415  GLPTTIGFHQRIMENPQFLQGNVSTSFV  442
```

```
RESULT 10
US-09-767-479-8
; Sequence 8, Application US/09767479
; Patent No. US20010036654A1
; GENERAL INFORMATION:
;   APPLICANT: Haselkorn, Robert
;              Gornicki, Piotr
;   TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
;                       METHODS OF USE
;   NUMBER OF SEQUENCES: 29
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Arnold, White & Durkee
;     STREET: P.O. Box 4433
;     CITY: Houston
;     STATE: Texas
;     COUNTRY: United States of America
;     ZIP: 77210
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/767,479
;     FILING DATE: 22-Jan-2001
;     CLASSIFICATION: Unknown
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/468,793
;     FILING DATE: <Unknown>
;     APPLICATION NUMBER: US SN 07/956,700
;     FILING DATE: 02-OCT-1992
;     APPLICATION NUMBER: PCT/US93/09340
;     FILING DATE: 30-SEP-1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Kitchell, Barbara S.
;     REGISTRATION NUMBER: 33,928
;     REFERENCE/DOCKET NUMBER: ARCD:152/KIT
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (512) 418-3000
;     TELEFAX: (713) 789-2679
;     TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 8:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 453 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-767-479-8

Query Match             17.1%;  Score 1005.5;  DB 9;  Length 453;
Best Local Similarity   47.1%;  Pred. No. 6.1e-72;
Matches  209;  Conservative  72;  Mismatches  158;  Indels  5;  Gaps  3;

Qy   29  FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY  88
            ||||||||||||  |  |  | :|  || :|
Db    3  FNKILIANRGEIALRILRTCEELGIGTIAVHSTVDRNALHVQLADEAVCIG-EAASSKSY  61

Qy   89  LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLIGDKSRAV  148
            |:||||||||||||||||||| |  :|||| | | |  | | | |||
Db   62  LNIPNIIAAALTRNASAIHPGYGFLAENARFAEICADHHLTFIGPSPDSIRAMGDKSTAK  121

Qy  149  TAAKKAGLPVLAESTP-SKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDELRKL  207
            ||||  | :|:| ::     :|||| |:||:|| ||||||:|||||||
Db  122  ETMQRVGVPTIPGSDGLLTDVDSAAKVAAEIGYPVMIKATAGGGGRGMRLVREPADLEKL  181

Qy  208  ATEASREAEAAFGDGSVVYERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVE  267
            |:| | |  | :|| ||:|||:|||||:|||||:||:|||
Db  182  FLAAQGEAFAAFGNPGLYLEKFIDRPRHVEFQILADAYGNVHLGERDCSIQRRHQKLLE  241
```

```
Qy  205 RKLATEASREAEAAFGDGSVVYERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQK 264
            ||  |||    :: ||| ::|:::|||:::||||||  :|: |||||:|||||:||||:||
Db  177 EQAFPRVISEATKAFGSAEVFLEKCIVNPKHIEAQILADSFGNTVHLFERDCSIQRRNQK 236

Qy  265 VVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEH 324
         :||||| ||:  | | || :||  ||| |||:|||||:|||  |:| |||||:|:|||||
Db  237 LIEIAPSPQLITPEQRAYIGDLAVRAAKAVGYENAGTVEFLLAD-GEYYFMEMNTRVQVEH 295

Qy  325 TVTEEVTEVDLVKAQMRLAAGATLKEIGLTQDKITHGAALQCRITTEDPNNGFRPDTGT 384
         |:||| | :|:|:| ||:||    |:||| | | : | ||| |||:||||:|| |:||
Db  296 TITEEITGIDVVREQIRIASGL--ELSVKQDDIVHRGYALQFRINAEDPKNNFLPSFGK 352

Qy  325 ITAYRSPGGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSG 444
         || | :||| ||  |:|| ||:| :||| | | : ||||:|   :||  ||| :||:|
Db  296 ITRYYAPGGPGVRTDTAIYTGYTIPPYYDSMCLKLIVWALIWEEALDRGLRALDDMRVQG 412

Qy  445 VATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQ 479
          : ||| :: :  :: | :: | ||  | :| |:|
Db  413 VKTTAPYYQEILRNPEFRSGQFNTSFVESHPELTQ 447
```

```
Qy  268 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 327
         ||   || : |  ||:   |  || |||:|||||:||||: | |:|||||||||||| | |
Db  242 EAPSPALSADLRQKMGDAAVKVAQIGYIGAGTVEFLVDATGNFYFMEMNTRIQVEHPVI 301

Qy  328 EEVTEVDLVKAQMRLAAGATLKEIGLTQDKITHGAALQCRITTEDPNNGFRPDTGTIITA 387
         |:: :|||| ||:|:| | ||:|||: :  |  | |: ||| |||| || |:|  | : |
Db  302 EMITGLDLIAEQIRIAQGEALR--FRQADIQLRGHAIECRINAEDPEYNFRPNPGRITG 358

Qy  388 YRSPGGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 447
         | :||| ||   ||   ::  ||||| |||  ||  ||  ||:  ||||  :||||  |
Db  359 YLPPGGPGVRDSHVYTDYEIPPYYDSLIGKLIVWGATREEAIARMQRALRECAITGLPT 418

Qy  448 NIGFLRALLREEDFTSKRIATGFI 471
          | |::: :| |:: | :::| |
Db  419 TLSFHQLMLQMPEFLRGELYTNFV 442
```

RESULT 11
US-09-815-242-5215
; Sequence 5215, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5215
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5215

Query Match   16.7%; Score 979.5; DB 9; Length 471;
Best Local Similarity 45.9%; Pred. No. 7.9e-70;
Matches 209; Conservative 74; Mismatches 157; Indels 15; Gaps 6;

```
Qy   30 KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL 89
         |||:|||||||||  | :||  |  : |:|| |:| :  : :|  | :| | |||| ||
Db    3 KKILIANRGEIAVIVRACAEMGIRSVAVYSEADRHALHVKRADEAHSIGAD--PLAGYL 60

Qy   90 DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTEVLDTGDKSRAVT 149
          : |::  | ::  | ::|||||||||| |||:|||.|| |||| :::  ||| | ::
Db   61 NPRALVNLAVESGCDALHPGYGFLSENAELAEICAERGIKFIGPSAQVIRRMGDKTEARR 120

Qy  150 AAKKAGLPVLAESTPSK--NIDDI---VKSAEGQTYPIFVKAVAGGGRGMRFVSSPDEL 204
         | :|||:|  ::: ||    |:|:    |:|  | | |::|||:||||||:| :: :: |
Db  121 SMIAAGVP---CTPGTEGNVADLAEALREAERIGYPVMLKATSGGGGRGIRRCNSREEL 176
```

RESULT 12
US-09-815-242-13885
; Sequence 13885, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13885
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-13885

Query Match   16.3%; Score 956; DB 9; Length 449;
Best Local Similarity 47.6%; Pred. No. 5.6e-68;
Matches 206; Conservative 61; Mismatches 158; Indels 8; Gaps 5;

```
Qy   31 KILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLD 90
         ||:|||||||| | | |:| ||  | ||   :  :  | : |:| | :| |  || || :
Db    4 KIVIANRGEIAIRILRACKELGIKTVAVSSADRDLKHVLLADETVCIGPAPS-VKSYLN 62

Qy   91 IDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDTGDKSRAVTA 150
          : |: |||  :|| | ||||||||||   |:  | ||  ||||  :  : | ||: | |
Db   63 IPAIISAAEITGAVAIHPGYGFLSENANFAEQVERSGFIFIGPKADTIRLMGDKVSAITA 122
```

```
Qy  151  AKKAGLPVLAESTPSKNIDDIVKSAEGQT--YPIFVKAVAGGGRGMRFVSSPDELRKLA  208
              ||||:| ::|||||   :||: ::|| ||||||| |||||||||||::    :: ||
Db  123  MKKAGVPTVPGSDGPLGDDMNANRAHAKRIGYPVIIKASGGGGRGMRFVVRSDAELAQSI  182

Qy  209  TEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEI  268
              :||| :|||:|::| ||::: ||||:||||:|||: |  ||:||||||:|||||||||:
Db  183  SMTKAEAKAAFSNDMVYMEKYLENPRHIEIQVLADGQGNAIYLAERDCSMQRRHQKVVEE  242

Qy  269  APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE  328
             ||| ||||||: || :|| : ||:|||||||:|||| ||::|:|||||:|||| ||||
Db  243  APAPGITPELRRYIGERCAKACVDIGYRGAGTFEFLF-ENGEFYFIEMNTRIQVEHPVTE  301

Qy  329  EVTEYDVLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAY  388
              :|| ||:|:| ||:|||   : ::| ||| :  ||:| :|||:|||:|| :|  |:|
Db  302  MITGVDLIKEQLRIAAG---QPLSTIQDEVVVRGHAVECRINAEDPNT-FLPSPGKITRF   357

Qy  389  RSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATN  448
              |:||  |||  |:  ||| ||:|||::|:  | | || ||:| :: ||:|::|  ||:|
Db  358  HAPGGFGVRWESHIYAGYTVPPYTDSMIGKLICYGENRDVAIARMKNALQELIIDGIKTN  417

Qy  449  IGFLRALLREDF  461
              | |||||:| :||
Db  418  IDLQTRIMNDEHF  430
```

RESULT 13
US-09-815-242-11160
; Sequence 11160, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
;   APPLICANT: Haselbeck, Robert
;   APPLICANT: Ohlsen, Kari L.
;   APPLICANT: Zyskind, Judith W.
;   APPLICANT: Wall, Daniel
;   APPLICANT: Trawick, John D.
;   APPLICANT: Carr, Grant J.
;   APPLICANT: Yamamoto, Robert T.
;   APPLICANT: Xu, H. Howard
;   TITLE OF INVENTION: Identification of Essential Genes in
    Prokaryotes
;   FILE REFERENCE: ELITRA.011A
;   CURRENT APPLICATION NUMBER: US/09/815,242
;   CURRENT FILING DATE: 2001-03-21
;   PRIOR APPLICATION NUMBER: 60/191,078
;   PRIOR FILING DATE: 2000-03-21
;   PRIOR APPLICATION NUMBER: 60/206,848
;   PRIOR FILING DATE: 2000-05-23
;   PRIOR APPLICATION NUMBER: 60/207,727
;   PRIOR FILING DATE: 2000-05-26
;   PRIOR APPLICATION NUMBER: 60/242,578
;   PRIOR FILING DATE: 2000-10-23
;   PRIOR APPLICATION NUMBER: 60/253,625
;   PRIOR FILING DATE: 2000-11-27
;   PRIOR APPLICATION NUMBER: 60/257,931
;   PRIOR FILING DATE: 2000-12-22
;   PRIOR APPLICATION NUMBER: 60/269,308
;   PRIOR FILING DATE: 2001-02-16
;   NUMBER OF SEQ ID NOS: 14110
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11160
;   LENGTH: 448
;   TYPE: PRT
;   ORGANISM: Haemophilus influenzae
US-09-815-242-11160

Query Match   16.2%; Score 950; DB 9; Pred. No. 1.7e-67;
Best Local Similarity 48.3%;
Matches 212; Conservative 57; Mismatches 152; Indels 18; Gaps 9;

```
Qy   30  KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL  89
              :|| ||||||:|    :: ||::: :||: :  :|  ::::::: | |: ||  ||::||
Db    3  EKVIANRGEIALRILRACKELGIKTVAVHSTADRDLKHVLLADETICIGPAPS-AKSYL   61
```

```
Qy   90  DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT  149
              :|: || |||||||:|||||:||||||||:   |||::||| | |:|||||:|||:|:|
Db   62  NIPAIIAAAEVTGADAIHPGYGFLSENADFAEQVERSGFTIGPTADVIRLMGDKVSAIK   121

Qy  150  AAKKAGLPVLAEST-PSKNIDDIVKS---AEGQTYPIFVKAVAGGGRGMRFVSSPDELR   205
              |:|||| |  ::|   |:|| | ||:     ::|:||||| |||||||||||:: |: ||
Db  122  AMKKAGVPCVPGSDGPVSN--DIAKNEIAKRIGYPIIIKASGGGGRGMRVVRSEDALE   179

Qy  206  KLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKV  265
              :  :||| :|||:|::| ||::: ||||:||||:|||: |  ||:||||||:||||||||
Db  180  ESIAMKAEAKAAFNNDMVYMEKYLENPRHVEIQVLADTHGNAVYLAERDCSMQRRHQKV  239

Qy  266  VEIAPAQHLDPELRDRI---CADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQV  322
              || ||| ||||||: :||   || :|| : ||:||||||:|||| ||::|:|||||:|||
Db  240  VEEAPAPGITEEVRRDIGSRCANA---CVEIGYRGAGTFEFLY-ENGEFYFIEMNTRIQV  295

Qy  323  EHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDT  382
              || ||| :| ||:|:| ||:|||:   : ::| ||| : :||:| :|||:|||:|| :
Db  296  EHPVTEMITGVDLVKEQLRIAAGL--PISFKQEDIKVKGHAMECRINAEDPKT-FLPSP   351

Qy  383  GTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTV  442
              | |:|: |:||  |||  |:  ||| ||:|||::|:  | | || ||:| :: ||:|::
Db  352  GKVNHLHSPGGLGVRWDSHVYGGYTVPPHYDSMIAKLITYGDTREVAIRRMQNALSETII  411

Qy  443  SGVATNIGFLRALLREEDF  461
              ::| ||| |||||:| :||
Db  412  DGIKTNIPLHELILEDENF  430
```

RESULT 14
US-10-156-761-14226
; Sequence 14226, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;   APPLICANT: OMURA, SATOSHI
;   APPLICANT: IKEDA, HARUO
;   APPLICANT: ISHIKAWA, JUN
;   APPLICANT: HORIKAWA, HIROSHI
;   APPLICANT: SHIBA, TADAYOSHI
;   APPLICANT: SAKAKI, YOSHIYUKI
;   APPLICANT: HATTORI, MASAHIRA
;   TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;   FILE REFERENCE: 249-262
;   CURRENT APPLICATION NUMBER: US/10/156,761
;   CURRENT FILING DATE: 2002-05-29
;   PRIOR APPLICATION NUMBER: JP 2001-204089
;   PRIOR FILING DATE: 2001-05-30
;   PRIOR APPLICATION NUMBER: JP 2001-272697
;   PRIOR FILING DATE: 2001-08-02
;   NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14226
;   LENGTH: 1171
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-14226

Query Match   16.0%;   Score 941; DB 15;   Length 1171;
Best Local Similarity 28.2%; Pred. No. 4.3e-66;
Matches 370; Conservative 146; Mismatches 464; Indels 334; Gaps 46;

```
Qy   29  FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVK-A  87
              : :|||||||||:| |   ::||:: : |||: :  :|  ||::::: | :|| :|||  |
Db    3  FDTLLVANRGEIAVRVIRTARELGLRTVAVYSDPDRSAPHVRLADEAVRLGP--APAKES  60

Qy   88  YLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRA  147
              |||  :|: ||| : :||| |||||||::|| |:||: :||.|.||||| |: :|| |:|
Db   61  YLDADLVLKAAKDTGAGAIHPGYGFLSEDAAFARRCEDAGIVFVGPTPEQLELFGAKHTA  120

Qy  148  VTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDELRKL  207
              :  |: || ||  ::: ::   :|  ::   :|| ||||:||||| || ::   ||:|
Db  121  RAAAEAAGVPLAPGTGLLASiDEALDAASRIGYPVMLKATGGGGIGMSACRSAAEELAES  180

Qy  208  ATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVE  267
```

Db   1114 VGFGDAVTAGQPLLTLEAMKMESRVPAPMAGVVTELLARPGDQVAAGTALVVLA 1167

RESULT 15
US-09-815-242-10330
; Sequence 10330, Application US/09815242
; Patent No. US2002006156 9A1
; GENERAL INFORMATION:
;   APPLICANT: Haselbeck, Robert
;   APPLICANT: Ohlsen, Kari L.
;   APPLICANT: Zyskind, Judith W.
;   APPLICANT: Wall, Daniel
;   APPLICANT: Trawick, John D.
;   APPLICANT: Carr, Grant J.
;   APPLICANT: Yamamoto, Robert T.
;   APPLICANT: Xu, H. Howard
;   TITLE OF INVENTION: Identification of Essential Genes in
;   TITLE OF INVENTION: Prokaryotes
;   FILE REFERENCE: ELITRA.011A
;   CURRENT APPLICATION NUMBER: US/09/815,242
;   CURRENT FILING DATE: 2001-03-21
;   PRIOR APPLICATION NUMBER: 60/191,078
;   PRIOR FILING DATE: 2000-03-21
;   PRIOR APPLICATION NUMBER: 60/206,848
;   PRIOR FILING DATE: 2000-05-23
;   PRIOR APPLICATION NUMBER: 60/207,727
;   PRIOR FILING DATE: 2000-05-26
;   PRIOR APPLICATION NUMBER: 60/242,578
;   PRIOR FILING DATE: 2000-10-23
;   PRIOR APPLICATION NUMBER: 60/253,625
;   PRIOR FILING DATE: 2000-11-27
;   PRIOR APPLICATION NUMBER: 60/257,931
;   PRIOR FILING DATE: 2000-12-22
;   PRIOR APPLICATION NUMBER: 60/269,308
;   PRIOR FILING DATE: 2001-02-16
;   NUMBER OF SEQ ID NOS: 14110
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10330
;   LENGTH: 449
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-09-815-242-10330

Query Match   16.0%; Score 940; DB 9; Length 449;
Best Local Similarity 46.8%; Pred. No. 1.1e-66;
Matches 203; Conservative 66; Mismatches 155; Indels 10; Gaps 6;

Qy   31 KILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLD 90
              ||:|||| | : || :||| |||:: || |:::::|||||:|: |: |||: |||
Db   4 KIVIANRGEIALRILRACKELGIKTVAVHSSADRDLKHVLLADETVCIGPAPS-VKSYLN 62

Qy   91 IDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTEVLDLTGDKSRAVTA 150
            ||:|||||:| ||||||||||||||| : | || ::| ||||||::||||:||||
Db   63 IPAIISAAEITGAVAIHPGYGFLSENANFAEQVERSGFIFIGPKAETIRLMGDKVSAIAA 122

Qy   151 AKKAGLPVLAESTPSKNIDDIVKS---AEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKL 207
            ||| ||||| ::: ||:|:||:::   :| | :|| |:|:|||||||||| ||:|| |
Db   123 MKKAGVPCVPGSDGPLG--DDMDKNRAIAKRIGYPVIIKASGGGGRGMRVVRGDAELAQS 181

Qy   208 ATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVE 267
            ||:: | ||||| ||||:||:| ||::|||| ||| | ||||||||||||||||||||||
Db   182 ISMTRAEAKAAFSNDMVYMEKYLENPRHVEIQVLADGQGNAIYLAERDCSMQRRHQKVVE 241

Qy   268 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 327
            ||||:||||||||||||||||||||||||||||||||| |:|||||||||||||||||||
Db   242 EAPAPGITPELRRYIGERCAKACVDIGYRGAGTFEFLF-ENGEFYFIEMNTRIQVEHPVT 300

Qy   328 EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 387
            || ||||||||||||||||||:||||||||| ||||:||||:|::||||||:|| :|| |
Db   301 EMITGVDLIKEQLRIAAG--- QPLSIKQEEVVHRGHAVECRINAEDPNT-FLPSPGKITR 356

Qy   388 YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 447
            | :||||||||||||:||||||:|::: | :: | : |:|| :|| :|: |:: |||

Db   357 FHAPGGFGVRWESHIYAGYTVPPYYDSMIGKLICYGENRDVAIARMKNALQELIIDGIKT 416

Qy   448 NIGFLRALLREEDF 461
          |: :: :|:|
Db   417 NVDLQIRIMNDENF 430

Search completed: September 24, 2003, 15:53:07
Job time : 44.7701 secs

```
                     GenCore version 5.1.6
                  Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:        September 24, 2003, 15:51:53 ; Search time 35.259 Seconds
                                        (without alignments)
                                        3155.707 Million cell updates/sec

Title:         US-09-974-973A-2
Perfect score: 5865
Sequence:      1 MTAITLGGLLKGIITILVST........RVVVPAATKVEGGDLIVVVS 1157

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:      283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters:      283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     PIR_76:*
               1: pir1:*
               2: pir2:*
               3: pir3:*
               4: pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES

Result              %
No.     Score     Match  Length DB  ID          Description
-----   ------    -----  ------ --  --          -----------
  1    3713.5     63.3    1127   2  D70671       pyruvate carboxyla
  2    2550.5     43.5    1178   2  A47255       pyruvate carboxyla
  3    2548.5     43.5    1178   2  JC4391       pyruvate carboxyla
  4    2535.5     43.2    1150   2  A83978       pyruvate carboxyla
  5    2530       43.1    1148   2  F69685       pyruvate carboxyla
  6    2528.5     43.1    1178   1  JC2460       pyruvate carboxyla
  7    2494       42.5    1146   2  AH1208       pyruvate carboxyla
  8    2491       42.5    1146   2  AC1565       pyruvate carboxyla
  9    2490.5     42.5    1144   2  D97227       pyruvate carboxyla
 10    2489       42.4    1150   2  G89881       pyruvate carboxyla
 11    2474.5     42.2    1174   2  AE2911       pyruvate carboxyla
 12    2474.5     42.2    1174   2  C97686       pyruvate carboxyla
 13    2465.5     42.0    1158   2  AE3285       pyruvate carboxyla
 14    2461       42.0    1175   2  T20346       pyruvate carboxyla
 15    2457       41.9    1178   1  QYBYP        pyruvate carboxyla
 16    2447       41.7    1185   2  T39734       pyruvate carboxyla
 17    2436       41.5    1180   2  S46094       pyruvate carboxyla
 18    2417.5     41.2    1195   2  T43735       pyruvate carboxyla
 19    2413       41.1    1137   2  E86708       pyruvate carboxyla
 20    2219.5     36.1     984   2  T44608       pyruvate carboxyla
 21    1079       18.4     501   2  D64453       biotin carboxylase
 22    1071.5     18.3     477   2  G70427       biotin carboxylase
 23    1040       17.7     472   2  A70432       biotin carboxylase
 24    1032.5     17.6     447   2  A53311       biotin carboxylase
 25    1032.5     17.6     447   1  AH1923       biotin carboxylase
 26    1020       17.4     506   2  D69277       biotin carboxylase
 27     996       17.0     491   2  A69123       biotin carboxylase
 28     979.5     16.7     471   2  G82966       probable biotin ca
 29     978.5     16.7     448   2  S74380       biotin carboxylase
 30     973.5     16.6    1095   2  B83471       probable pyruvate
 31     963.5     16.4    1078   2  D87647       hypothetical prote
 32     962.5     16.4     447   2  B97338       biotin carboxylase
 33     956       16.3     449   2  AI0912       biotin carboxylase
 34     955       16.3     667   2  F98286       hypothetical prote
 35     955       16.3     677   2  AC2997       hypothetical prote
 36     954.5     16.3     444   2  C70444       biotin carboxylase
 37     954.5     16.3     455   2  B86722       biotin carboxylase
 38     950       16.2     449   2  AD0445       biotin carboxylase
 39     950       16.2     448   1  F64105       biotin carboxylase
 40     948.5     16.2     539   2  T07093       acetyl-CoA carboxy
 41     944.5     16.1     457   2  H71553       probable biotin ca
 42     940       16.0     449   2  JS0632       acetyl-CoA carboxy
 43     939       16.0     449   2  D85990       hypothetical prote
 44     938       16.0     444   2  T44813       biotin carboxylase
 45     938       16.0     449   2  H91144       hypothetical prote
```

ALIGNMENTS

```
RESULT 1
D70671
pyruvate carboxylase (EC 6.4.1.1) - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 11-Jan-2002
C;Accession: D70671; S73055
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
;Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70671
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1127 <COL>
A;Cross-references: GB:Z83018; GB:AL123456; NID:g3261671; PIDN:CAB05410.1; PID:g16948
A;Experimental source: strain H37Rv
R;Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, September 1994
A;Description: Mycobacterium tuberculosis cosmid tbc2.
A;Reference number: S73053
A;Accession: S73055
A;Molecule type: DNA
A;Residues: 1-353,'TRAGSARCDPPAVPVSAWTAAPTWRRNQPVLRLHAGQADLS',396-1115,'EWRAETCWWW' <
A;Cross-references: EMBL:U00024; NID:g560506; PIDN:AAA50948.1; PID:g560527
C;Genetics:
A;Gene: pca; PYC
A;Start codon: GTG
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindi
C;Keywords: biotin binding; ligase; mitochondrion
F;4-457/Domain: biotin carboxylase homology <BCH>
F;1055-1127/Domain: lipoyl/biotin-binding homology <LPB>
F;1093/Binding site: biotin (Lys) (covalent) #status predicted

  Query Match             63.3%;  Score 3713.5;  DB 2;  Length 1127;
  Best Local Similarity   64.4%;  Pred. No. 2.4e-196;
  Matches  730;   Conservative 153;   Mismatches 239;   Indels  11;   Gaps   6;

Qy      29  FKKIIVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY  88
                                                                                                    
Db       2  FSKVLVANRGEIAIRAFRAAYELGVGTVAVYPYEDRNSQHRLKADESYQIGDIGHPVHAY  61

Qy      89  LDIDELIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 148
            |:||||:   |:|:   ||||||||||||||| |:|:|:|  ||:|||:|||||| :|||
Db      62  LSVDEIVATARRAGADAIYPGYGFLSENPDLAAACAAAGISFVGPSAEVLELAGNKSRAI 121

Qy     149  TAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDELRKLA 208
            |||||||| | | |:|| |: :: ||::: :|||:|||||||||||||||||||||||||
Db     122  AAARRAGLPVLMSSAPSASVDELLSVAAGMFPFLFVKAVAGGGRGMRRVGDIAALPEAI 181
```

A;Title: Adipose pyruvate carboxylase: amino acid sequence and domain structure deduc
A;Reference number: A47255; MUID:93189578; PMID:8445588
A;Accession: A47255
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1178 <ZHA>
A;Cross-references: GB:L09192; NID:g293743; PIDN:AAA39737.1; PID:g293744
A;Experimental source: 3T3-L1 adipocytes
A;Note: sequence extracted from NCBI backbone (NCBIN:126874, NCBIP:126875)
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindi
C;Keywords: biotin binding; ligase; mitochondrion
F;1-20/Domain: transit peptide (mitochondrion) #status predicted <TRP>
F;21-178/Product: pyruvate carboxylase #status predicted <MAT>
F;39-494/Domain: biotin carboxylase homology <BCH>
F;1105-1178/Domain: lipoyl/biotin-binding homology <LPB>
F;1144/Binding site: biotin (Lys) (covalent) #status predicted

Query Match   43.5%;   Score 2550.5;   DB 1; Length 1178;
Best Local Similarity 46.1%;   Pred. No. 2.6e-132;
Matches 543;   Conservative 192;   Mismatches 407;   Indels 35;   Gaps 13;

```
Qy      7 GGLLLKGIITLVSTHTSSTLPA-------------FKKILVANRGEIAVRAFRAALETGAA  54
            |||||    ||||        ||             |||||||||||||||||| |
Db      9 GGLRLLGV------RRSSAPVASPNVRRLEYKPIKKVMVANRGEIAIRVFRACTELGIR  62

Qy     55 TVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLDIDELIGAAKKVKADAIYPGYGFLS 114
            |  |||  ||| |||| |    ||||  ||||||||||| ||||||| ||||||||||||
Db     63 TVAVYSEQDTGQMHRQKADEAYLIGRGLAPVQAYLHIPDIIKVAKENGVDAVHPGYGFLS 122

Qy    115 ENAQLARECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVL-AESTPSKNIDJIVK 173
            :||  |||| | ||| ||| ||||  |:||  |||    ||||| |    :   ||:||
Db    123 ERADFAQACQDAGVRFIGPSPEVVRKMGDKVEARAIAIAAGVPVVPGTDSPISSLHEAHE 182

Qy    174 SAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKLATEASREEAAFGDGSVVVERAVINP 233
             |||| | || |||||||||||||||||  |||  | |  ||    ||   |   ||
Db    183 FSNTFGFPIIFKAYGGGGRGMRVVHSYEELENYTRAYSEALAAFGNGALFVEKFIEKP 242

Qy    234 QHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVETAPAQHLDPELRDRICADAVKFCRSI 293
            :|| | |||||   ||| |||||  :||  ||| ||  |  :::||  :::
Db    243 RHIEVQILGDQYGNILHLYERDCSIQRRHQKVVEIAPATHLDPQLRSRLTSDSVKLAKQV 302

Qy    294 GYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKEIGL 353
            ||  | || |||||||||||||||| | ||| ||| |||||||| |  | | |  |||
Db    303 GYENAGTVEFLVDKHGKHYFIEVNSRLQVEHTVEEITDVDLLVHAQIHVSEGRSLPDIGL 362

Qy    354 TQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQL-GGEITAHF 412
            |  | ||||||||||||||| | ||||||| ||||||| ||||     |  |||| |
Db    363 RQENIRINGCAIQCRVTTEDPARSFQPDTGRIEVFRSGEGMGIRLDNASAFQGAVISPHY 422

Qy    413 DSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIG 472
            |:||| | ||||  |||||| | ||||| |:|| ||  ||||||  |||  |||||||||
Db    423 DSLLVKVIAHGKDHPTAATKMSRALAEFRVGKTNIPFLQNVLNNQQFLAGTVDTQFID 482

Qy    473 DHPHLLQAPPADDEQGRILDYLADVTVNKPGVGRVPKDVA-APID-KLPNIKDLPLPRGSR 530
            | ||  ||| |||  ||| |||  ||||| | |    :| | ||  || | |||  |
Db    483 ENPELFQLRPAQNRAQKLLHYLGHVMVNGPTTPIPVNVSPSVPDAVPVVPIGPPAGFR 542

Qy    531 DRLKQLGPAAFARDLREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELL 590
            :: | | || | ||  ||| ||:| :||||||||||||||  :||||| | ||  || |:
Db    543 DILLREGPEGFARAVRNHQGLLLMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFNKLF 602

Qy    591 SVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAF 650
            |:     ||  ||||| || :  :|| | | |||| ||||||| |||||||||||||| |
Db    603 SMENWGGATFDVAMRFLYECPWRRLQELRELIPNIPFQMLLRGANAVGYTNYPDNVVFKF 662

Qy    651 VKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLD 710
            :||  ||| :|:|::|:  :|  ||  |:|    :| | ||      |::|  :||| |:
Db    663 CEVAKENGMDVFRVFDSLNYLPNMLLGMEARGSAG-GVVEAAISYTGDVADPSRTKYSLE 721

Qy    711 YYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVTHDTAGGQLA 769
            |||  |||:||| || |||||||||||| |  | |:  ||  |:|::|||||||||| |
Db    722 YYMGLAEELVRAGTHILCIKDMAGLLKPAACTMLVSSLRDRFPDLPLHIHTHDTSGAGVA 781
```

RESULT 2
A47255
Pyruvate carboxylase (EC 6.4.1.1) precursor [similarity] - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Feb-2002
C;Accession: A47255
R;Zhang, J.; Xia, W.L.; Brew, K.; Ahmad, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 1766-1770, 1993

```
Qy    209 TEASREAEAFGDGSVYVERAVINPOHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEI 268
            :||:|||| :| ::||||||||||:|||||||||||||||||||||||||||||||
Db    182 EAASREASAFGDPTVYLEQAVINPRHIEVQILADNLGDVIHLYERDCSVQRRHQKVIEL 241

Qy    269 APAQHLDPELRDRICADAVKFCRSIGYQOGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE 328
            ||:|:||| ||||::|  ||:|   ||  :| ||||:|||||:|||||||| | |||| |
Db    242 APAPHLDAELRYKMCVDAVAFARHIGYSCAGTVEFLLDERGEYVFIEMNPRVQVEHTVTE 301

Qy    329 EVTEVDLVKAQMRLAAGATLKELGLITQDKIKTHGAALQCRITTEDPNNGFRPDTGITAY 388
            | ||:|||||:|||||||||:|  ||  :|  |||||||||||||||||||||||||
Db    302 EITDVDLVASQLRIAAGETLEQLGRQEDIAPHGAALQCRITTEDPANGFRPDTGRISAL 361

Qy    389 RSPGGAGVRLDGAAQLGGEITAHEDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATN 448
            | :|||||||| ::| |  |:  | ||||:|||| |||  |||||||| | |  ||||
Db    362 RTAGGAGVRLDGSTNLGAEISPYFDSMLVKLTCRGRDLPTAVSRARRAIAEFRIRGVSTN 421

Qy    449 IGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKPHGVRPK 508
            ||||||||||||||:|||||||  | :|||||| |||  ||| |||:||||||  ||
Db    422 IPFLQAVLDDPDFRAGRVTSFIDERPQLLTARASADRGTKILNFLADVTVNNPYGSRPS 481

Qy    509 DVAAPIDKLPNIKDLPL----PRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDAHQS 564
              |  ::|||:|| ||    |||| || ||| | ||| ||| | ||||| |||||||||
Db    482 TI-YPDDKLP---DLDLRAAPPAGSKQRLVKLGPEGFARWLRESAAVGVTDTTFRDAHQS 537

Qy    565 LLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPN 624
            |||||||| |     |:| || |  || |||||||||||||||||||||  |:|||||||
Db    538 LLATRVRTSGLSRVAPYLARTMPQLLSVECWGGATYDVALRFLKEDPWERLATLRAAMPN 597

Qy    625 VNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVLET 684
            : :|||||||||||||| |||  ||||:||||||||||||||||||||||||||||:|||
Db    598 ICLQMLLRGRNTVGYTPYPEIVTSAFVQEATATGIDIFRIFDALNNIESMRPAIDAVRET 657

Qy    685 NTAVAEVAMAYSGDLSDPNEKLYTLDYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKL 744
            |:| |||||| ||| :| | ||||| ||| | ||:| |||:||||||||||||||  |
Db    658 GSAIAEVAMCYTGDLTDPGEQLYTLDYLKLAEQIVDAGAHVLAIKDMAGLLRPPAAQRL 717

Qy    745 VTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAA 804
            :||||| || |||||||||||| || |||| ||||   |:| || | ||||| |||||
Db    718 VSALRSRFDLPVHLTHDTPGGQLASYVAAWHAGADAVDGAAAPLAGTTSQPALSSIVAA 777

Qy    805 FAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFFESGTPGVYRHEIPGGQLSNLRAQA 864
            || | ||| | |   ||||||||   | | :||||: ||||||||||||||||||| |
Db    778 AAHTEYDIGLSLSAVCALEPYWEALRKVYAPFESGLPGPTGVYHEIPGGQLSNLRQQA 837

Qy    865 TALGLADRFELIEDNYAAVNEMLGRPTKVIPSSKVVGDLALHLVGAGVDPADFAADPQKY 924
            ||||| |||| ||    | | | ||||  |   ||||| || |:||| |:|| ||| | |
Db    838 IALGLGDRFEIEEAYAGADRVLGRLVKVTPTSKVVGDLALALVGAGVSADEFASDPARF 897

Qy    925 DIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDADDSKERR 984
            ||| | || ||||| || ||||| ||| ||  |||| ||||||||  :|| || |:||
Db    898 GIPESVLGFLRGELGDPPGGWPEPLRTAALAGRGAAR-PTAQLAADDEIALSSVGAK-RQ 955

Qy    985 NSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDA 1044
            :|||||||||||||||  ||||| | |||||  |||:|||| | |  | |||:|||::|
Db    956 ATLNRLLFSPTKEFNEHRPAYGDTSQLSANQFFYGLRQGEEHRVKL-ERGVELLIGLEA 1014

Qy   1045 ISEPDDKGMRNVANVNGQIRPMRYRDRSVESVIATABEKADSSNKGHVAAPFAGVVTVTV 1104
            ||||| ||| ||:| || ||| |||| || |:| |||   |:|| |  ||||| |||||
Db   1015 ISEPDERGMRTVMCIINGQLRPVLVRDRSTASAVPAAEKADRGNPGHIAAPFAGVVTVGV 1074

Qy   1105 AEGDEVKAGDAVAIIEAMKMEATTTASVDGKIERVVPAATKVEGGDLIVVS 1157
            ||| ||| | || |||||||||| |||  | ||:| |||||||||||:|||
Db   1075 CVGERVGAGQTIATIEAMKMEAMTTAPVAGVTVERVAVSDTAQVEGGDLLVVS 1127
```

```
F;1105-1178/Domain: lipoyl/biotin-binding homology <LPB>
F;1144/Binding site: biotin (Lys) (covalent) #status experimental

Query Match          43.5%;     Score 2548.5; DB 2;  Length 1178;
Best Local Similarity 46.3%;     Pred. No. 3.3e-132;
Matches  545; Conservative 188; Mismatches 409; Indels 35; Gaps 13;

Qy    7 GGLLLKGITLVSTHTSSLPA------------FKKILVANRGEIAVRAFRAALETGAA  54
         ||||||:|||||||||||         ||| :|||||||||| |:|| |  ||
Db    9 GGLRLLGV-----RRSSTAPVASPNVRRLEYKPIKKVMVANRGEIAIRVFRACTELGIR  62

Qy   55 TVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLDIDELIGAAKKVKADAIYPGYGFLS 114
         ||| |  :  | :||  | ||| || |:|| ||||  | ||: ||:  |: |:|||||||
Db   63 TVAVYSEQTPGQMHRQKADEAYLIGRGLAPVQAYLHIPDIIKVAKENGVDAVHPGYGFLS 122

Qy  115 ENAQLARECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVL-AESTPSKNIDDIVK 173
         | | :|:||::|| ||||| ||| |  ||| |||  |:||:||| ||   | | : :|::
Db  123 ERADFAQACQDAGVRFIGPSPEVVRKMGDKVEARAIAIAAGVPVVPGTNSPINSLHEAHE 182

Qy  174 SAEQGTYPIFVKAVAGGGGRGMRFVSSPDELRKLATEASREAEAAFGDGSVYVERAVINP 233
          :  | :||:.||| ||||||||.| | |||:   |: ||:|:|||:|:::||:.||:.|
Db  183 FSNTYGFPIIFKAAYGGGGRGMRVVHSYEELEENYTRAYSEALAAFGNGALFVEKFIEKP 242

Qy  234 QHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSI 293
         :||||||||| .| ::|||||||.||||||||||||| :||||:| .|: :| ||:::::
Db  243 RHIEVQILGDYGNILHLYERDCSIQRRHQKVVEIAPATHLDPQLRSRLTSDSVKLAKQV 302

Qy  294 GYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGL 353
         || .||||||||| .| |.||| | |:||||||||| |.||||.|| : | | :| :|||
Db  303 GYENAGTVEFLVDKHGKHYFIEVNSRLQVEHTVTEEITDVDLVHAQIHVSEGRSLPDLGL 362

Qy  354 TQDIKTHGAALQCRITTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQL-GGEITAHF 412
          |:|: . .| ||| |||| :..|:|||| | .:||  | .:||| ||: |||.: :::
Db  363 RQENIRINGCAIQCRVTTEDPARSFQPDTGRIEVFRSGEGMGIRLDNASAFQGAVISPHY 422

Qy  413 DSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIG 472
         || |||:   |.| :|| ::::|||||| |.|| ||| |: .:|:::.|:.|| .:.||:
Db  423 DSLLVKVIAHGKDHPTAATKMSRALAEFRVRGVKTNIPFLQNVLNNQQFLAGIVDTQFID 482

Qy  473 DHPHLLQAPPADDEQGRILDYLADYTVNKPHGVRPKDVA-APIDKL-PNIKDLPLPRGSR 530
         ::| |:: |  .::: |:| ||  : ||  ||  | ::  |||.:| |:|  || | |.|
Db  483 ENPELFQLRPAQNRAQKLLHYLGHVMVNGPTTPIPVKVSSPVDPIVPVVPIGPPAGFR 542

Qy  531 DRLKQLGPAAFARDLREQDALAVDTTFRDAHQSLLATVRSFALKPAEEAVAKLTPELL 590
         | |::|||. |||:|||:  |:|||||||||||||||||. | |:||:|| |: :.||
Db  543 DILLREGPEGFARAVRNHQCLLLMDTTFRDAHQSLLATVRVTHDLKKIAPYVAHNFNNLF 602

Qy  591 SVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVIQMLLRGRNTVGYTPYPDSVCRAF 650
         |:| |||||:||||||| | || ::.||||::|.:  ||||| | |||| ||| |  :|
Db  603 SIENWGGATFDVAMRFLYECPWRRIQELRELIPNIPFQMLLRGANAVGYTNYPDNVVFKF 662

Qy  651 VKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLD 710
          .||.: |:|:|||||:||:: ::   ::. :|   :||: :  | ||: || .|:|.|:
Db  663 CEVAKENGMDVFRIFDSLNYLPNMLLGMEAAGSAG-GVVEAATSYTGDVADPSRTKYSLE 721

Qy  711 YYILKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVTHDTAGGQLA 769
         || : :|||:|:.| |||.|||||||| |||.:.||:.:| .| ||| |:|||:: .  |
Db  722 YYMGLAEELVRAGTHILCIKDMAGLLKPAACTMLVSSLRDRFPDLPLHIHTHDTSGSGVA 781

Qy  770 TYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAV 829
          :  ||||||| || |  .:||:|||||:.|:||     | :  ||  |  |:  ||| |
Db  782 AMLACAQAGADVVDVAVDSMSGMTSQPSMGALVACTKGTPLDTEVPLERVFDYSEYWEGA 841

Qy  830 RGLYLPFESGTPGPTGR--VYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEML 887
         |||| .|:. |   :.   |:.:|||||| :||  :| :.:| :.|: : :.:.||||:|
Db  842 RGLYAAFDCTATMKSGNSDVYENEIPGGQYTNLHFQAHSMGLGSKFKEVKKAYVEANQML 901

Qy  888 GRPTKVTPSSKVVGDLALHLVGAGVDPDAFAADPQKYDIPDSVIAFLRGELGNPPGWPE 947
         |    ||.|||:|||||::::  | . :|| |  :    .:||: ||:|::| ||||.||
Db  902 GDLIKVTFSSKIVGDLAQFMVQNGLSRAEAEAQAEELSFPRSVVEFLQGYIGIPHGGFPE 961

Qy  948 PLRTRAL-----EGRSEGKAPLTEVPEEQAHLDA-DDSKERRNSLNRLLFPKPTEEFL 1000
          :||::|     ||| .| :|| . |||:  ::  :. |:::|||||||||| |||||
Db  962 PFRSKVLKDLPRIEGRPGASLPLNLKELEKDLIDRHGEEVTPEDVLSAAMYPDVFAQFK 1021

Qy 1001 EHRRRFGNTSALDDREEFYGLVEGRETLIRLRPDVRFPLLVRLDAISEPDDKGMRNVVANV 1060
         ::       | ||::   ||:|||:: : ::    |  : :|:: :.| :.| :|  .
Db 1022 DFTATFGPLDSLNTRLFLQGPKIAEEFEVELERGKI-LHIKALAVSDLNRAGQRQVFFEL 1080

Qy 1061 NGQIRPMRVRDRSVESVTATAEKADSNKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAII 1119
         |||:|.: |:| ::.| ||:||::. .: ||:||  | |:.:.||.||:|| ||   :|
Db 1081 NGQLRSILVKDTQAMKEMHFHPKALKDVKGQIGAPMPGKVIDIKVVAAGDKVAKGQPLCVL 1140

Qy 1120 EAMKMRATITASVDGKIERVVVPAAATKVEGGDLIVVV 1156
         .||||:.: |:. .|.| :||  | : .:.|||::::
Db 1141 SAMKMETVVTSPMEGTIRKVHVTKDMTLEGDDLILEI 1177


RESULT 3
JC4391

Pyruvate carboxylase (EC 6.4.1.1) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Jan-1996 #sequence_revision 25-Apr-1997 #text_change 11-Jan-2002
C;Accession: S68252; S72393; JC4391; S06440
R;Jitrapakdee, S.; Booker, G.W.; Cassady, A.I.; Wallace, J.C.
  Biochem. J. 316, 631-637, 1996
A;Title: Cloning, sequencing and expression of rat liver Pyruvate carboxylase.
A;Reference number: S68252; MUID:96257760; PMID:8687410
A;Accession: S68252
A;Molecule type: mRNA
A;Residues: 1-1178 <JIT1>
A;Cross-references: EMBL:U36585; NID:g1040973; PIDN:AAC52668.1; PID:g1040974
A;Accession: S72393
A;Molecule type: protein
A;Residues: 489-505 <JIT2>
A;Experimental source: liver
R;Lehn, D.A.; Moran, S.M.; MacDonald, M.J.
  Gene 165, 331-332, 1995
A;Title: The sequence of the rat pyruvate carboxylase-encoding cDNA.
A;Reference number: JC4391; MUID:96096545; PMID:8522203
A;Accession: JC4391
A;Molecule type: mRNA
A;Residues: 1-221,'P',223-865,'D',867-976,'G',978-1178 <LEH>
A;Cross-references: GB:U32314; NID:g929987; PIDN:AAA96256.1; PID:g929988
A;Experimental source: liver
R;Thampy, K.G.; Huang, W.Y.; Wakil, S.J.
  Arch. Biochem. Biophys. 266, 270-276, 1988
A;Title: A rapid purification method for rat liver pyruvate carboxylase and amino acids
A;Reference number: S06440; MUID:89024676; PMID:3178228
A;Accession: S06440
A;Molecule type: protein
A;Residues: 'SG',23-25,'PL',28-29,'LL',32-34,'P',1134,'A',1136-1137,1139-1152,'T',1154-1...
C;Comment: This enzyme is located in the mitochondrial matrix and catalyzes the conversi...
   intermediates that exit the mitochondrion for consumption in various pathways.
C;Genetics:
A;Gene: pc
A;Genome: nuclear
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding
C;Keywords: biotin binding; gluconeogenesis; homotetramer; ligase; mitochondrion
F;1-20/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F;21-1178/Product: pyruvate carboxylase #status predicted <MAT>
F;39-494/Domain: biotin carboxylase homology <BCH>
```

```
Db   962 PFRSKVLKDLPRIEGRPGASLPPLNLKELEKDLIDRHGEEVTPEDVLSAAMYPDVFAQFK 1021
Qy  1001 EHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANV 1060
Db  1022 DFTATFGPLDSLNTRLFLQGPKIAEEFEVELERGKT-LHIKALAVSDLNRAGORQVFFEL 1080
Qy  1061 NGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAII 1119
Db  1081 NGQLRSILVKDTQAMKEMHFIPKALKDVKQGIGAPMPGKVIDVKVAAGAKVVKGQPLCVL 1140
Qy  1120 EAMKMEATITASVDGKIERVVVPAATKVEGGDLIVVV 1156
Db  1141 SAMKMETVVISPMEGTIRKVHVHVTKDMTLEGDDLILEI 1177
```

RESULT 4
A83978
pyruvate carboxylase pycA [imported] - Bacillus halodurans
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: A83978
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujì, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: A83978
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1150 <STO>
A;Cross-references: GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BAB06344.1; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
A;Gene: PycA
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

```
Query Match           43.2%;   Score 2535.5;   DB 2;      Length 1150;
Best Local Similarity 46.9%;   Pred. No. 1.7e-131;
Matches  540;     Conservative 185;   Mismatches 399;   Indels 27;   Gaps 13;

Qy   26 LPAFKKILVANRGEIAVRAFRAAALETGAAATVAIYPREDRGSFHRSFASRAVRIGTEGSSPV 85
Db    4 LKNIKKVLVANRGEIAIRIFRACTELHIRTVAIYSKEDTGAYHRYKADEAYLVGEGKKPI 63
Qy   86 KAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKS 145
Db   64 EAYLDIEGIIEIAKRHGVDAIHPGYGFLSENIEPAKRCHEEGITFIGPELEHLVMFGDKV 123
Qy  146 RAVTAAKKAGLPVLAEST-PSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDEL 204
Db  124 QAREQAIKANLPVIPGSDGPVSSLEDVKAFADKHGYPFIIKAALGGGGRGMRIVRSENDV 183
Qy  205 RKLATEASREAEAAFGDGSVYVERAVINPHIEVOILGDRTGEVVHLYERDCSLQRRHQK 264
Db  184 QESYERAKSEAKAAFGNDEVYVEKFIENPKHIEVQILADKHGNTLHLYERDCSVQRRHQK 243
Qy  265 VVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEH 324
Db  244 VVEVAPSVSLSEDVRERICQAAVQLAENVNYVNAGTVEFLVDREGNFYFIEVNPRIQVEH 303
Qy  325 TVTEEVTEVDLVKAQMRLLAAGATL-KELGL-TQDKIKTHGAALQCRITTEDPNNGFRPD 381
Db  304 TITEMVTGIDIVQSOLFIADGEHLHGDRLGIPKQEEIVCHGYAIQSRVTTEDPSNGFLPD 363
Qy  382 TGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEF 440
Db  364 TGRINAYRSGGGFGVRLDAGNFQGAVITPYYDSLLVKVSTWALTFEGAAAKMLRNLREF 423
Qy  441 TVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDLLADVTVN 500
Db  424 RIRGIKTNIAFLENVVQHQFLSGEYNTSFIDQTPELFVFPKKRKDRGTKMLSFTGETIVN 483
Qy  501 KPHGVRPKDVAAPIDK--LPNIK-DLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 557
```

```
Db  484 GYPGLE-KTKKPVFDKPPVPKLKLSEP-IPDGTKQILDQHGPEGLAKWVKEQKHVLLTDTT 542
Qy  558 FRDAHQSLLATRVRSFALKPAAFAVAKLTPELLSVEAWGGATYDVAMRPLFEDPWDRLDE 617
Db  543 FRDAHQSLLATRVRTHDLKQIAEPTARLLPNLFSMEMWGGATFDVAMRFLHEDPWERLLI 602
Qy  618 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 677
Db  603 LRKKAPNVLFQMLLRASNAVGYKNYPDNLIREFVDKSANAGIDVFRIFDSLNWWEGMKLA 662
Qy  678 IDAVLEINTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR 737
Db  663 IEAVGEAN-KIAEATICYTGDILDSSRPKYDLAYYKKIAKELEAAGAHILGIKDMAGLLK 721
Qy  738 PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 797
Db  722 PEAAYQLVAELKDTVIIPVHLHTHDTSGNGIFTYARAIEAGVDIVDVAVSSMAGLTSQPS 781
Qy  798 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 857
Db  782 ANSLYYALADSERQPNVNITALEQLAEFWEETRKFYAGFESGMNAPHTEVYEHEMPGGQY 841
Qy  858 SNLRAQATALGLADRFLIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 917
Db  842 SNLQQQAKAVGLGHRWNEVKKMYRTVNDMFGDVVKVTPSSKVVGDMALYMVQNDLTEEV 901
Qy  918 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDA 977
Db  902 YENGHGKLDFPDSVVEFFEGQLGQPYQGFPKKIQEIILKGRK----PITNRPGENMEPIQF 957
Qy  978 DDSKER-RNSLNR----------LLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRE 1026
Db  958 EAIKEELYNKLDRQVTSHDILSYALYPKVFMEFERFRQTFGDVSVLLDTPTFYGLRPGEE 1017
Qy 1027 TLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADS 1086
Db 1018 IEVEIEQGKT-LIVKFISLSKPQDDGNRIVYFELNGQPREVLIKDQSVKTSIISRPKADK 1076
Qy 1087 SNKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVPAAT 1145
Db 1077 SNPNHIGASMPGTVVKALVEKGDKVKQGDHLMITEAMKMETTVQAPFDGEVVALHVKDGD 1136
Qy 1146 KVEGGDLIVVV 1156
Db 1137 AIQTGDLLIEV 1147
```

RESULT 5
F69685
pyruvate carboxylase (EC 6.4.1.1) pycA [similarity] - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 18-Aug-2000
C;Accession: F69685
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: F69685
A;Status: preliminary: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1148 <KUN>

A;Cross-references: GB:Z99111; GB:AL009126; NID:g2633699; PIDN:CAB13359.1; PID:g2633857
A;Experimental source: strain 168
C;Genetics:
A;Gene: PycA
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding
C;Keywords: ligase
F;8-465/Domain: biotin carboxylase homology <BCH>
F;1073-1146/Domain: lipoyl/biotin-binding homology <LPB>
F;1112/Binding site: biotin (Lys) (covalent) #status predicted

Query Match   43.1%;   Score 2530;   DB 2;   Length 1148;
Best Local Similarity   46.0%;   Pred. No. 3.3e-131;
Matches 533;   Conservative 188;   Mismatches 390;   Indels 48;   Gaps 13;

```
Qy     28  AFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKA    87
               :||::||||||:|| || || ||   :|:::|    :|| |||: :: ||:|:||:|
Db      5  SIQKVLVANRGEIAIRIFRACTELNIRTVAVYSKEDSGSYHRYKADEAYLVGEGKKPIDA    64

Qy     88  YLDIDELIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRA   147
               |||| ||| :||:|||||| ||:|||||| ||| | ||| ||||||| | || |: |
Db     65  YLDIEGIIDIAKRNKVDAIHPGYGFLSENIHFARRCEEEGIVFIGPKSEHLDMFGDKVKA   124

Qy    148  VTAAKKAGLPVLAEST-PSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRK   206
               ||:::||:||:|||  | ::|:| :| : |||::|||| || :|||||| :: |:|
Db    125  REQAEKAGIPVIPGSDGPAETLEAVEQFGQANGYPIIIKASLGGGGRGMRIVRSESEVKE   184

Qy    207  LATEASREAEAAFGDGSVYVFERAVINPQHIEVQILGDRTGEVHLYERDCSLQRRHQKVV   266
               |||||  ||||||| :||: ||||::| :|||:||||| || |||||||:|:|||||||
Db    185  AYERAKSEAKAAFGNDEVYVEKLIENPKHIEVQVIGDKQGNVHLFERDCSVQRRHQKVI   244

Qy    267  EIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTV   326
               | || | || || || ||||:|||:|:|||||||:||:|||:|||:|| ::||||||||:
Db    245  EVAPSVSLSPELRDQICEEAAVALAKNVNYINAGTVEFLV-ANNEFYFIEVNPRVQVEHTI   303

Qy    327  TEEVTEVDLVKAQMRLAAGATL--KELGLITQDK-IKTHGAALQCRITTEDPNNGFRPDTG   383
               |:|:|:||::||:| ||:   : |:||:   ::| :|| || :|: | |||||||| ||||
Db    304  TEMITGVDIVQTQILVAQGHSLHSKKVNIPEQKDIFTIGYAIQSRVTTEDPQNDFMPDTG   363

Qy    384  TITAYRSPGGAGVRLD-GAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTV   442
               |||| || ||||||| || | |  || | |:: :|| ||| |: ||:|:|||:::|:|
Db    364  KIMAYRSGGGFGVRLDTGNSFQGAVITPYYDSLLVKLSTWALTFEQQAAAKMVRNLQEFRI   423

Qy    443  SGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVN--   500
               ||:|:||:|| |||:|||:| | |  || ||  || |||  | :||| |  |:|:|:
Db    424  RGIKTNIPFLENVAKHEKFLTGQYDTSFIDTTPELFNFPKQKDRGTKMLTYIGNVTVNGF   483

Qy    501  ------KPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAV   553
                     |  | : |  || ||::|: ||| ::|  ||||:|| |||::|| |:|:|||
Db    484  PGIGKKEKPAFDKPLGVKVDVDQQP-------ARGTKQILDEKGAEGLANWVKEQKSVLL   536

Qy    554  TDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLEDDPWD   613
               ||||||||||||||||:|||||||| || ||||||||||||||||||||| ||||:||:
Db    537  TDTTFRDAHQSLLATRIRSRHDLKKIANPTAALWPELFSMEMWGGATFDVAYRFLKEDPWK   596

Qy    614  RLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQ   673
               |||:|||:||||:|::| |||| ||:|||||| | :|||| :|||||:|||||||||:|
Db    597  RLEDLRKEVPNTLFQMLLRSSNAVGYTNYPDNVIKEFVKQSAQSGIDVFRIFDSLNWVKKG   656

Qy    674  MRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMA   733
               ||| |||| || ||||:|| |||||||||:||||||:|| |:||| |:|||||||||||
Db    657  MTLAIDAVRDTG-KVAEAAICYTGDILDKNRTKYDLAYYTSMAKELEAAGAHILGIKDMA   715

Qy    734  GLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTT   793
               ||||:|| | | || |:|||:||:||||:|:|:||| | |||:|||||:|| |:| | ||
Db    716  GLLKPQAAYELVSALKETIDIPVHLHTHDTSGNGIYMYAKAVEAGVDIIDVAVSSMAGLT   775

Qy    794  SQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPGTGRVYRHEIP   853
               |||| | | ||  ||| |:| |:|| |  |||  |: | || | || |:|: |||  |
Db    776  SQPSASGFYHAMEGNDRRPEMNVQGVELLSQYWESVRKYYSEFESGMKSPHTEIYEHEMP   835

Qy    854  GGQLSNLRAQATALGLADREFLIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVD   913
               ||| |||: |  ||| | ||::| ||   :: :|||  || :|||||||| |:|| :|:|
Db    836  GGQYSNLQQAKGVGLGDRWNEVKEMYRRVNDMFGDIVKVTPSSKVVGDMALYMVQNNLT   895
```

RESULT 6
JC2460
pyruvate carboxylase (EC 6.4.1.1) precursor - human
N;Alternate names: pyruvate:carbon dioxide ligase (ADP-forming)
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 18-Aug-2000 #text_change 01-Feb-2002
C;Accession: G01933; JC2460; B27883; S01469
R;Walker, M.E.; Jitrapakdee, S.; Val, D.L.; Wallace, J.C.
submitted to the EMBL Data Library, July 1995
A;Reference number: H00708
A;Accession: G01933
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1178 <WAL>
A;Cross-references: EMBL:U30891; NID:g1101028; PIDN:g1101028; PIDN:AAA82937.1; PID:g1101029
R;MacKay, N.; Rigat, B.; Douglas, C.; Chen, H.S.; Robinson, B.H.
Biochem. Biophys. Res. Commun. 202, 1009-1014, 1994
A;Title: cDNA cloning of human kidney pyruvate carboxylase.
A;Reference number: JC2460; MUID:94324922; PMID:8048912
A;Accession: JC2460
A;Molecule type: mRNA
A;Residues: 1-224,'WP',227-351,'A',353-384,'PT',387-485,'DV',488-637,'R',639-728,'A',
A;Cross-references: GB:S72370; NID:g632807; PIDN:AAB31500.1; PID:g632808
R;Lamhonwah, A.M.; Quan, F.; Gravel, R.A.
Arch. Biochem. Biophys. 254, 631-636, 1987
A;Title: Sequence homology around the biotin-binding site of human propionyl-CoA carb
A;Reference number: A27883; MUID:87212051; PMID:3555348
A;Accession: B27883
A;Molecule type: mRNA
A;Residues: 1083-1178 <LAM>
A;Cross-references: GB:M26122; NID:g189657; PIDN:AAA36423.1; PID:g387003
R;Freytag, S.O.; Collier, K.J.
J. Biol. Chem. 259, 12831-12837, 1984
A;Title: Molecular cloning of a cDNA for human pyruvate carboxylase. Structural relat
A;Reference number: S01469; MUID:85030380; PMID:6548474
A;Accession: S01469
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1135-1178 <FRE>
A;Cross-references: EMBL:K02282; NID:g189655; PIDN:AAA60033.1; PID:g189656
C;Genetics:
A;Gene: GDB:PC
A;Cross-references: GDB:119472; OMIM:266150
A;Map position: 11q11-11q13.1
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindi
C;Keywords: biotin binding; gluconeogenesis; ligase; mitochondrion
F;1-20/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F;21-1178/Product: pyruvate carboxylase (mitochondrion) #status predicted <MAT>
F;39-494/Domain: biotin carboxylase homology <BCH>
F;1105-1178/Domain: lipoyl/biotin-binding homology <LPB>
F;1144/Binding site: biotin (Lys) (covalent) #status predicted

```
Qy    914  PADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEE---   970
               |: | |:  :|:|: |: |: |:||||:|| :::|| ||    ||:|:||||::
Db    896  EKDVYEKGESLDFPDSVVELFKGNIGQPHGGFPEKLQKLILKGQE---PITVRPGELLE   951

Qy    971  ---------EQAHLPADDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFF  1018
                        :|| |||:|| || :::|| ||  |:| | |:||::
Db    952  PVSFEAIKQEFKEQHNLEISD----QDAVAYALYPKVFTDYVKTESYGDISVLDTPTFF  1007

Qy   1019  YGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNQGIRPMVRDRSVESVT  1078
               | ||| | |: : |  |   |||   :  || |:|| |:|| | :|:||||:| :::|:
Db   1008  YGMTLGEIEVEIERGKT-LIVKLISIGEPQPATRVVVFELNGQPREVVIKDESIKSSV  1066

Qy   1079  ATAEKADSSNKGHVAAPFAGVVTVTVAE-GDEVKAGDAVAIIEAMKMERATITASVDGKIE  1137
               || :|| :  ||:|  |||| |:|:  | |: |||| ::   |||| | | | :|||| |
Db   1067  QERLKADRINPSHIAASMPGTVIKVLAEAGTKVNKGDHLMINEAMKMETTVQAPFSGTIK  1126

Qy   1138  RVVVPAATKVEGGDLIVVV  1156
               :|| || ||:| ||::::
Db   1127  QVHVKNGEPIQTGDLLLEI  1145
```

```
Query Match          43.1%;  Score 2528.5;  DB 1;  Length 1178;
Best Local Similarity 45.9%;  Pred. No. 4.2e-131;
Matches  540;  Conservative 187;  Mismatches 415;  Indels  35;  Gaps  13;

Qy      7 GGLLLKGIITLVSTHTSSTLPA----------FKKILVANRGEIAVRAFRAALETGAA  54
          ||||| ||:    ||| ||                ||||||||||||||||||||||| 
Db      9 GGIRLLGI------------RRTSTAPAASPNVRRLEYKPIKKVMVANRGEIAIRVFRACTELIGIR  62

Qy     55 TVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLS  114
          |||||  || :|| | ||| | |:|:| ||||||||||||:|| ||:|:|||||||||||
Db     63 TVAIYSEQDTGQMHRQKADEAYLIGRGLAPVQAYLHIPDIIVKAKENNVDAVHPGYGFLS  122

Qy    115 ENAQLARECAENGITFIGPTPEVLDLTGDKSRAVTRAAKKAGLPVL-AESTPSKNIDDIVK  173
          :||||:||||||||||||||||| |:| | ||||:|  | |||||  |: |:|:|||||||
Db    123 ERADFAQACQDAGVRFIGPSPEVVRKMGDKVEKARAIAIAAGVPVVPGTDAPITSLHEAHE  182

Qy    174 SAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKLATEASREAEAAFGDGSVVVERAVINP  233
          |  ||||:||||||||||||||||||  :|||| |   ||:|:|||| ||||||||:|||
Db    183 FSNTYGFPIIFKAAYGGGGRGMRVVHSYEELEENYTRAYSEALAAFGNGALFVEKFIEKP  242

Qy    234 QHIEVQILGDRTGEVVHLYERDCSLQRRHQKVEIAPAQHLDPELRDRICADAVKFCRSI  293
          :|||||| || | || ||||||||||||||||| |:||||||||||||||||||||||:
Db    243 RHIEVQILGDQYGNILHLYERDCSIQRRHQKVEIAPAAHLDPQLRLRLTSDSVKLAKQV  302

Qy    294 GYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGL  353
          |||||||||||||:||| ||||||||||||||||||:||||||||:||||||||||||||
Db    303 GYENAGTVEFLVDRHGKHYFIEVNSRIQVEHTVTEEITDVDLVHAQIHVSEGRSLPDLGL  362

Qy    354 TQDKITHGAALQCRITTEDPNNGFRPDTGITAYRSPGGAGVRLDGAAQL-GGEITAHF  412
          ||| | |||||||:| : :|| ||  |||| | | ||||||||||:|| |||:| |||
Db    363 RQENIRINGCAIQCRVTTEDPARSFQPDTGRIEVFRSGEGMGIRLDNASAFQGAVISPHY  422

Qy    413 DSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIG  472
          |||| || |  ||| |||||||:|||||| || |:|| | | ||:|||:|||:|  |:|
Db    423 DSLLVKVIAHGKDHPTAATKMSRALAFFRVRGVKTNIAFLQNVLNNQQFLAGTVDTQFID  482

Qy    473 DHPHLLQAPPADDEQGRILDYLADVTVNKPHGVRP-KDVAAPIDK-LPNIKDLPLPRGSR  530
          :|| |  |||  ::||||| |:|||| ||||| ||    ||| |  |||:||||||||:|
Db    483 ENPELFQLRPAQNRAQKLLHYLGHVMVNGPTTIPIPVKASPSPTDPVVPAVPIGPPAGFR  542

Qy    531 DRLKQLGPAAFARDLREQDALAVTDTITFRDAHQSLLATRVRSFALKPAAEAVAKLTPELL  590
          ||||:||||| |||||:|||| ||| |||||||||||||||||||||| :| |:||| ||
Db    543 DILLREGPEGFARAVRNHPGLLMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFSKLF  602

Qy    591 SVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAF  650
          | || ||||||||||| | ||  ||:|:  :||| ||:||||||:| ||||||||:|||:
Db    603 SMENWGGATFDVAMRFLYECPWRRLQELRELIPNIPFQMLLRGANAVGYTNYPDNVVFKF  662

Qy    651 VKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLD  710
          |||||| |::: ::||  | :| ||   | |:||||| || :  ||||||||| |:| ||
Db    663 CEVAKENGMDVFRVFDSLNYLPNMLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQ  721

Qy    711 YYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVHTHDTAGGQLA  769
          |||||:|||::|||:|  |||| | | |||: |::  | |:    |:||:|||||:||:|
Db    722 YYMGLAEELVRAGTHILCIKDMAGLLKPTACTMLVSSLRDRFPDLPLHIHTHDTSGAGVA  781

Qy    770 TYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSEAVSDLEPYWEAV  829
          ||:|| ||||| |  || || | : | | |||| :| :|:| ||| ||  ||| |||| 
Db    782 AMLACAQAGADVVDVAADSMSGMTSQPSMGALVACTRGTPLDTEVPMERVFDYSEYWEGA  841

Qy    830 RGLYLPFESGTPGPTGR--VYRHEIPGGQLSNLRAQATALGLADREFLIEDNYAAVNEML  887
          ||||  |: | || |:    | |: ||||:|:||||:  | |||||:||||||:|| | |
Db    842 RGLYAAFDCTATMKSGNSDVENEIPGGQYINLHFQAHSMGLGSKFKEVKKAYVEANQML  901

Qy    888 GRPTKVTPSSKVVGDLALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPE  947
          ||| ||||||||||||| || ||| :|:  |  ||:||:|  || |||||  ||| ||  
Db    902 GDLIKVTPSSKIVGDLAQFMVQNGLSRAEAEQAEELSFPRSVVEFLQGYIGVPHGGFPE  961

Qy    948 PLRTRAL------EGRSEGKAPLTEVPEEEQAHLDA-DDSKERRNSLNRLLFKPKPTEEFL  1000
          ||:| |      |||| | |||:|||  ::|| ||  |:|:   | |||:|||| | |  |
Db    962 PFRSKVLKDLPRVEGRPGASLPPLDLQALEKELVDRHGEEVTPEDVLSAAMYPDVPAHFK  1021

Qy   1001 EHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLIVRLDAISEPDDKGMRNVVANV  1060
          ||||:||||||| |||||| |||||| |||| :: :|:| |||   :|||||||||||:
Db   1022 DFTATFGPLDSLNTRLFLQGPKIAEEFEVELERGKT-LHIKALAVSDLNRAGQRQVFFEL  1080

Qy   1061 NGQIRPMRVRDRSVESVIATAEKADSSNKGHVAAPFAG-VVTVIVAEGDEVKAGDAVAII  1119
          || :|:   || || :|  |    : :|| ||| |  |  | | : | ||||| ||::|:
Db   1081 NGQLRSILVKDTQAMKEMHFHPKALKDVKGQIGAPMPGKVIDIKVVAGAKVAKGQPLCVL  1140

Qy   1120 EAMKMEATITASVDGKIERVVPAATKVEGGDLIVVV  1156
          |||||||| | |:|  ||||||||||||| ||:||
Db   1141 SAMKMETVVTSPMEGTVRKVHVTKDMTLEGDDLILEI  1177
```

RESULT 7

AH1208

pyruvate carboxylase homolog pyca [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AH1208
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Bercie, P.; Bloec
 .; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A:Title: Comparative genomics of Listeria species.
A:Reference number: ABI077; MUID:21537279; PMID:11679669
A:Accession: AH1208
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1146 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAC99150.1; PID:gi16410474; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: PyCA
C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindi

```
Query Match          42.5%;  Score 2494;  DB 2;  Length 1146;
Best Local Similarity 45.7%;  Pred. No. 3.2e-129;
Matches  525;  Conservative 186;  Mismatches 408;  Indels  30;  Gaps  11;

Qy     30 KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL  89
          |||||||||||| | |||||||||| |||||||||||||||||||||||||||| |||| 
Db      5 KKVLVANRGEIAIRVMRACTELKIKTVAIYSQEDTGSFHRYKSDEAYLVGAGKKPIDAYL  64

Qy     90 DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT  149
          ||| || ||||||||:||||||||||||| | |  | ||||||| ||| :||:||||  
Db     65 DIENIIEIAKESGADAIHPGYGFLSENIEFARRCEQEGIIFVGPKSKHLDMFGDKIKAKE  124

Qy    150 AAKKAGLPVLAEST-PSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKLA  208
          | |||||||| |:   ||||||:||||||||:|||||| ||||||||||:|||||| :|
Db    125 QALLADIPVIPGSNGPVAGIKEVEEFGEKNGYPLMIKASLGGGGRGMRVVESKEHVKESF  184

Qy    209 TEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEI  268
          |:||||||| ||| ||:|||||:||:||||:|||| |||:||||||||||||||||||:|
Db    185 ERASSEAKAAFGNDEVYYEKCVMNPKHIEVQILGDTHGNIVHLFERDCSIQRRHQKVVEV  244

Qy    269 APAQHLDPELRDRICADAVKFCRSIGTQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE  328
          || ||||| ||| |||||||||||||||| ||||| ||||| || |||:||||||  :|
Db    245 APCNAITSELRNRICDAAVKLMKNVDYINAGTVEFLV-EGDDYFIEVNPRVQVEHTITE  303

Qy    329 EVTEVDLVKAQMRLAAGATLKELGLT---QDKIKTHGAALQCRITTEDPNNGFRPDTGTI  385
          ||||||||||| ||| |||::||||    |:||:|||||||  |: ||| | || ||: |
Db    304 MITGIDIVQSQLFIADGYALHDQLVAIPKQEDIHIHGSAIQSRITTEDPLNNFMPDTGRV  363

Qy    386 TAYRSPGGAGVRLD-GAAQLGGEITAHPDSMLVKMTCRGSDFETAVARAQRALAEFTVSG  444
          | |||:|||| ||| ||| || :|  || |:|||  | || |||| ||| |:|:|||:||
Db    364 DIYRSTGGFGVRLDAGNGFQGTVVTPFYDSLLVKLCTWGMTFEQATRKMRRNLIEFRIRG  423

Qy    445 VATNIGELRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKPHG  504
          | |||  | | |:||||:|| | || |  |  |   |||  ::|||||||||||||||| 
Db    424 VKTNIPFLLNVVRHPDFASGNYNTSFIDTTPELFKFPHIRDRGTKTLRYIGNVTVNGPFG  483
```

```
Query Match        42.5%;  Score 2491;  DB 2;  Length 1146;
Best Local Similarity   45.9%;  Pred. No. 4.7e-129;
Matches  527;  Conservative 183;  Mismatches 409;  Indels  30;  Gaps  12;

Qy     30 KKILVANRGEIAYRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL 89
          ||:|||||||||||||||||||| : | ||||||||| :|| |||
Db      5 KKVLVANRGEIAIRVMRACTELKIKTVAIYSQEDTGSFHRYKSDEAYLVGAGKKPIDAYL 64

Qy     90 DIDELIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLITGDKSRAVT 149
          |||:|| | ||: || | |||||||||||::::||||||| |||
Db     65 DIENIIEIAKESGADAIHPGYGFLSENIEFARRCEQEGIIFVGPKSKHLDMFGDKIKAKE 124

Qy    150 AAKKAGLPVIAEST-PSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKLA 208
          ||| ||| : ::| ::|||| | : |||||: :::
Db    125 QALLADIPVIPGSDGPVAGIKEVEEFGEKNGYPLMIKASLGGGGRGMRVVESKEHVKESF 184

Qy    209 TEASREARAAFGDGSVYYERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEI 268
          |||||||||||||||||:||||||||||||||:||||||||:||||||||||:||||||
Db    185 ERASSEAKAAFGNDEVYVEKCVMNPKHIEVQILGDTHGNIVHLFERDCSIQRRHQKVVEV 244

Qy    269 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE 328
          || || |||||||:|| : :|||| |||||||| ||| || |:||||||| |||: |||
Db    245 APCNAITSELRNRICDAAVKLMKNVDYINAGTVEFLV-EGDNFYFIEVNPRVQVEHTIIE 303

Qy    329 EVTEVDLVKAQMRLAAGATIKELGLT---QDKIKTHGAALQCRITTEDPNNGFRPDTGTI 385
          : | |:|| :||| ||| ||| |: |||::|||||||| |||| | | || |||
Db    304 MITGIDIVQSQLFIADGVALHDQLVAIPKQEDIHIHGSAIQSRITTEDPLNNFMPDTGRV 363

Qy    386 TAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSG 444
          | |||||| |||| ||| || |||:|| :||||| |:||:|||| | |||| | |||
Db    364 DTYRSTGFGVRLDAGNGFQGTVVTPFYDSLLVKLCTWGMTFEQATRKMRRNLIEFRIRG 423

Qy    445 VATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKPHG 504
          ||| | ||||||||||:|| |:|| || ||| ||:| |||||||: || ||||||:|||
Db    424 VKTNIPFLLNVVHPDFASGNYNTSFIDTTPELFKFPHIRDRGTKTLRYIGNVTVNGFPG 483

Qy    505 VRPKDVAAPIDKLPNIKDLP----LPRGSRDRLKQLGPAAFARDLREQDALAVTDITFRD 560
          |||||||||||||||| |: ||||||||:||| |||||| |||||||||| |||||||
Db    484 IKHRD--KPVYAEPRLPKIPYGSQISPGTKQILDAKGPEGVVDWVKKQEEVLLTDTTLRD 541

Qy    561 AHQSLLATRVRSFALKPAAEAVAKLIPELLSVEAWGGATYDVAMRFLFEDPWDRLDELRE 620
          |||||||||||||||||||||||||||| ||| |||||| |:| | :|| |: :|| |
Db    542 AHQSLLATRVRSKDIFQIADAMAHLLPNMFSFEMWGGATFDVAYRFLNEDPWVRLETLRK 601

Qy    621 AMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSOMRPAIDA 680
          |:|:| | ||||| | ||||||||| |||||:|:||||| || ||: |||:|||||
Db    602 QIPNVMFQMLLRGANAVGYKNYPDNVIREFVKQSAQSGVDVFRVFDSLNWIKGMEVSIDA 661

Qy    681 VLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAA 740
          |:| | |:||| | ||| :| :| ||||||||:||||: ||||| |:||||||||:|||
Db    662 VREAG-KVVEAICYTGDIDDTRTKYTIDYYKDMAKELVAQGTHILGIKDMAGLLKPQA 720

Qy    741 VTKLIVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSA 800
          | ||:|||||||||||| | |: ||||:||| || || |||| || :||:||:||||||
Db    721 AYRLIGELKDTVDPIHLHTHDTSGNGIYTYAAAVSAGVDIVDVASSAMSGATSQPSMTG 780

Qy    801 IVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFFESGTPGTGRVYRHEIPGGQLSNL 860
          |||||:|||||||:|| |:|:|| ::| ::|| |||| :|||||: : ||||||||:||
Db    781 LYYGLVNGNRQTNLDAQNSQIINHYWEDVRHYKDFDNALNSPQTEVVIHEMPGGQYINL 840

Qy    861 RAQATALGLADREFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAAD 920
          |||| ||||| ||:|||:| ||| || |:| |||||||||| ||| | :|| | || |
Db    841 QQQAIAVGLGDRWDEVKEMYTVNQMFGDIVKVTPSSKVVGILALFMVQNELTEEDVYEK 900

Qy    921 PQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLITEVP-------EEE 971
          | || |||||| || |||||||| | : | || || ||| | |||:||
Db    901 GDTIDFPDSVIEFFMGEIGQPYGGFPEKLQKLVLKGRT---PLADRPGALMEPVNFAEV 956

Qy    972 QAHLDADDSKE- RRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLI 1029
          :|:|| ||: |:||||| ||| ::| ||| || | |:| :|| | ::| |||||
Db    957 KAELKEKMGYEPSEKDVISYILYPKVFLDYQEMISKYGDVTVLDTPFYKGMRLGEFIEV 1016

Qy   1030 RLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNK 1089
```

RESULT 8
AC1565
pyruvate carboxylase homolog pycA [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AC1565
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AC1565
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1146 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC96291.1; PID:g16413519; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: pycA
C;Superfamily: Pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

---

```
Qy    505 VRPKDVAAPIDKLPNIKDLP----LPRGSRDRLKQLGPAAFARDLREQDALAVTDITFRD 560
          |||||||||||||||| |: ||||||||:||| |||||| |||||||||| |||||||
Db    484 IKHRD--KPVYAEPRLPKIPYGSQISPGTKQILDAKGPEGVVDWVKKQEEVLLTDTTLRD 541

Qy    561 AHQSLLATRVRSFALKPAAEAVAKLIPELLSVEAWGGATYDVAMRFLFEDPWDRLDELRE 620
          |||||||||||||||||||||||||||| ||| |||||| |:| | :|| |: :|| |
Db    542 AHQSLLATRVRSKDIFQIDDTRTKYTIDYYKDMAKELVAQGTHLGIKDMAGLLKPQA 720

Qy    621 AMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSOMRPAIDA 680
          |:|:| | ||||| | ||||||||| |||||:|:||||| || ||: |||:|||||
Db    602 QIPNVMFQMLLRGANAVGYKNYPDNVIREFVKQSAQSGVDVFRVFDSLNWIKGMEVSIDA 661

Qy    681 VLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAA 740
          |:| | |:||| | ||| :| :| ||||||||:||||: ||||| |:||||||||:|||
Db    662 VREAG-KVVEAICYTGDIDDTRTKYTIDYYKDMAKELVAQGTHILGIKDMAGLLKPQA 720

Qy    741 VTKLIVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSA 800
          | ||:|||||||||||| | |: ||||:||| || || |||| || :||:||:||||||
Db    721 AYRLIGELKDTVDPIHLHTHDTSGNGIYTYAAAVSAGVDIVDVASSAMSGATSQPSMTG 780

Qy    801 IVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFFESGTPGTGRVYRHEIPGGQLSNL 860
          |||||:|||||||:|| |:|:|| ::| ::|| |||| :|||||: : ||||||||:||
Db    781 LYYGLVNGNRQTNLDAQNSQIINHYWEDVRHYKDFDNALNSPQTEVVIHEMPGGQYTNL 840

Qy    861 RAQATALGLADREFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAAD 920
          |||| ||||| ||:|||:| ||| || |:| |||||||||| ||| | :|| | || |
Db    841 QQQAIAVGLGDRWDEVKEMYTVNQMFGDIVKVTPSSKVVGILALFMVQNELSEEDVYEK 900

Qy    921 PQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLITEVP-------EEE 971
          | || |||||| || |||||||| | : | || || ||| | |||:||
Db    901 GDTIDFPDSVIEFFMGEIGQPYGGFPEKLQKLVLKGRT---PLADRPGALMEPVNFAEV 956

Qy    972 QAHLDADDSKE-RRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLI 1029
          :|:|| ||: |:||||| ||| ::| ||| || | |:| :|| | ::| |||||
Db    957 KAELKEKMGYEPSEKDVISYILYPKVFLDYQEMISKYGDVTVLDTPFYKGMRLGEFIEV 1016

Qy   1030 RLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNK 1089
          |||||||||||| |:| : ||||||||||||||||||||||||||:|||| || ||||||
Db   1017 ELEKGKI-LLIKLNSIGEPIADGTRVIYFELNGQPREINIQDMNVQSTVIARRKIDTTNP 1075

Qy   1090 GHVAAPFAG-VTVTVAEGDEVKAGDAVAIIEAMKMEATTASVDGKIERVVVPAATKVE 1148
          | ||| ||| ||||||| || ||| |: |||||:| |:|||| || |:||: :||:
Db   1076 EHVGATMGSVIQVVVVKGDSVKKGDPLLITEAMKMETTIQAPFDGEVSSTYVSDGDTIE 1135

Qy   1149 GGDLIVVS 1157
          |||:: :
Db   1136 SGDLLEVN 1144
```

```
Qy    562  HQSLLATRVRSFALKPAAEAVAKITPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREA  621
           ||||:|||:: || : |:| :| :  :  ||| ||  |:|: |:||  ::|: : |:|
Db    540  HQSLMATFLRTVDMVKIAKAESVLAKDLFSMEMWGGATFDTAYRFLKESPWERLERLRKR  599

Qy    622  MPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAV  681
           || :|| ::|:| | | : : :|||: |  :|::|  |  |  :| :|:|  :: :|
Db    600  VPNVLFQMLLRGANAVGYKNYPDNVIREFIKQSSKSGIDVFRIFDSLNWVKGMEVAIDEV  659

Qy    682  LETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAAV  741
           |  :  |  ||: || : ||  |: ::|:|| ||  :| |||||| || :|||:||| ||
Db    660  LNQG-KVAEACMCYTGDILDINRDKYILNYYYNLAKEIEKSGAHILGKDMSALLKPYAA  718

Qy    742  TKLVTALRREFDLPVHVHTHDTAGGQLATYAAQAGADAVDGASAPLSGTTSQPSLSAI  801
           ||:| |:: :|  :|:|||:||| ||| ||: |:|  ||:|  | | :| ::: | :|:
Db    719  LKLIRALKNEISIPIHLTHDTTGNGVATVLMAAHAGVDIADTAFNSMSGLTSQPALNSV  778

Qy    802  VAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSNLR  861
           |||  |:||| | |: : |:|::||: |  |    :|| | |:: | :| :||| |:|:|
Db    779  VAALKNTDRDTKMDIGDLQKISDYWSTVRPVYSKFESGLKAVSAEIYKYEIPGGQYSNLK  838

Qy    862  AQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAADP  921
           |:||:|  ||:|| | ::|: :|  ||:|||| |||::|||| :|:|| | ||: ||:
Db    839  PQVESFGLGHRFEQVKEMYREVNIMLGDIVKVTPSSKMVGDLAIFMVQNELTSENILEKA  898

Qy    922  QKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDADSK  981
           |  |:|  | : |  :| |:     |     : |:  |   | :  | |::  :|  ||:
Db    899  KDMPFPDSVVSYFKGMMGQPKGGFPKELQKIVL---KDEFAITCRPGELLPDEDFDKIR  954

Qy    982  ERRNSLNRL------------LFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGREFLIR  1030
                 : |:|             | ||  | ::||||  | :|: | ||| :| |::| ::
Db    955  VRLKKENKIEPTDKDVISYALYPDVFEEYLKYKNEYGDLSRMGSDVFFHGLAEGEISELE  1014

Qy    1031 LPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMVRDRSVESVTATAEK--ADSS  1087
           | || : |||  ::| :||||| ::::|:|:| | |:: :|: ::| : | ||  ||||
Db    1015 IAEGKT-LVVQLLHIGKLDKQGNRTLVFEVNGNRREIKIKDKVSSTKSEIVEEIVIADSS  1073

Qy    1088 NKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVVPAATK  1146
           |||::|:    ||:||:||||| :|:| |:||||||:|||||:|  ||  |||: |||
Db    1074 NKKEIGASIPGNVVKVFVFVKPGDKVKKGDSLMVIEAMKMETNVSVSEDGTVGGIFVKEGDQ  1133

Qy    1147 VEGGDLIV  1154
           ||||| ||
Db    1134 VQSGQLIV  1141
```

RESULT 9

D97227

Pyruvate carboxylase, PYKA [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C;Accession: D97227
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, H.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: D97227
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1144 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK80607.1; PID:g15025689; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC2660
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

Query Match   42.5%; Score 2490.5; DB 2; Length 1144;
Best Local Similarity   45.2%; Pred. No. 5e-129;
Matches 519; Conservative 204; Mismatches 392; Indels 33; Gaps 14;

```
Qy    29   FKKTIVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY  88
           || ::|||||||| |||  |||:|| ||||||| |  ||| | |:::  :|| |||| |
Db    5    FKRVLVANRGEIAIRIFRACHELGIRTVAIYSEEDKLALFRTKADESYLIGQNKGPVDAY  64

Qy    89   LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITIGPTPEVLDLTGDKSRAV  148
           | ||:|:   | | : |:| ||||||| || ::| :|| | |||  |   || |: :
Db    65   LNIDEIIINLALKKGVDAIHPGYGFLSENSEFSRRCTEAGIEFIGPTGDMMDKLGDKINSK  124

Qy    149  TAAKKAGLPVL-AESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKL  207
           ||||:|| |:| |:|: ||: |: :  :  : || ||||  ||||||| |  |  ||||:|
Db    125  LAAKAAGVKTIPGVEKPIETEQQAIEFARTCGYPVMVKAAAGGGGRGMRIVEKEEDLIAA  184

Qy    208  ATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVE  267
           :| : | |::|| |   |::| || ||:||||  :  : :| :|||||||: ||:||:|:
Db    185  CRSAKSEAKKAFGIEDIFIEKYLEGPKHIEVQVLGDKYGNIVHLYERDCSVQRRHQKVIE  244

Qy    268  IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT  327
           |||:  ||||||:|::  ||:|  |||:|| ||||| |:|:|| |||||||| |||| |
Db    245  LTPAVSMSEEKRLEICEDALKIARSIGYRSAGTLEFLLDKHGNHYFIEMNPRVQVEHTIT  304

Qy    328  EEVEVDLVKAQMRLAAGATLK-ELGL-TQDKIKTHGAALQCRITTEDPNNGFRPDTGT  384
           |||||||| |: |:|||: ||  |  : ||| |:| |||  |||:|: |: | ||:|| |
Db    305  EMVIGIDIVQSQILLAEGYKLNSPEVGINSQEDIHVNGYAIQCRITTEDPSNSFAPDTGK  364

Qy    385  ITAYRSPGGAGVRLDGA-AQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVS  443
           | |:|   ||  |:|||  ::|| :  |:|  |:|: ||::|| ||:|| :|:| :||:
Db    365  IDVRTGSGFGIRLDGGNGFTGAVISPYDSLLVKSTWSRTFEDAIRKAIRAIKETYIS  424

Qy    444  GVATNIGFLRALLREEDTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKPH  503
           || |:|| |  |::: | :|| | ||:|| |  ::| ||:| : |||:||: |:|:|
Db    425  GVKTNIDFLNVLNHETFRKGLCDTNFIANNPELFEITPRIDTELRVLKFIGEKVVNETH  484

Qy    504  GVRPKDVAAPIDKLP--NIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVIDTTFRDA  561
           ||| :|:  ||:||   |:| || : | :||  :|  || || || :||||:||||||:|
Db    485  GHK---IEVDVPSVPKYEIKE-PL-RGTKQILDEKGPKGLVEWIKDQDKLLLTDTTMRDA  539

Qy    30   KKILVANRGEIAVRAFRAALETGAATVAIYPREDKGRSFHRSFASEAVRIGTEGSPVKAYL  89
           |:||||||||||||||||||||||||||||||||| ::||:| ||||||||||||||||||
```

RESULT 10

G89881

pyruvate carboxylase [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: G89881
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: G89881
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1150 <KUR>
A;Cross-references: GB:BA000018; PID:g137009115; PIDN:BAB42211.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: pycA
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindi

Query Match   42.4%; Score 2489; DB 2; Length 1150;
Best Local Similarity   45.2%; Pred. No. 6.1e-129;
Matches 527; Conservative 193; Mismatches 383; Indels 62; Gaps 14;

```
Qy    30   KKILVANRGEIAVRAFRAALETGAATVAIYPREDKGRSFHRSFASEAVRIGTEGSPVKAYL  89
           |:|||||||||||||||||||||||||||||||| ::||:| ||||||||||||||||||
```

```
Qy   1132  VDGKIERVVVPAATKVEGGDLIVVV  1156
           |||:  | :::|  |  ::||  |
Db   1121  FDGVIKQVTVNNGDTIATGDLLIEI  1145
```

RESULT 11
AE2911

Pyruvate carboxylase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AE2911
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1174 <KUR>
A:Cross-references: GB:AE008688; PIDN:AAL43707.1; PID:g17741236; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: PycA
A:Map position: circular chromosome
C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindi

Query Match   42.2%;   Score 2474.5;   DB 2;   Length 1174;
Best Local Similarity   46.5%;   Pred. No. 3.9e-128;
Matches 540;   Conservative 174;   Mismatches 406;   Indels 41;   Gaps 19;
```

```
Qy    25  TLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIG------  79
                                 :|  ||  : | |  ::|         :   |  :
Db    20  TVLKISKILVANRSEIAIRVFRAANELGIKTVAIWAEEDKLSLHRFKADESYQVGRGPHL  79

Qy    80  -TEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVL  138
           |    |:||  : |    | :    || || ||||||||||||||    |  |  | ||
Db    80  AKDMGPTFSYLSEEVIRVAKLSGADAIHPGYGLLSESSPEFVEACNKAGITFIGPTPDTM  139

Qy   139  DLTGDKSRAVTAAKKAGLPVLAESTP-SKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRF  197
           | |||:|||  ||  |   ||  : |  |||:|||  |||||   ||||||||||||  |
Db   140  RQLGNKVAARNLAISVDVPVVPATNPLPDDIAEVERMAEEIGPVMLKASWGGGGRGMRA  199

Qy   198  VSSPDELRKLATEASREAEAAPGDGSVYVERAVINPQHIEVQILGDRTGEVHLYERDCS  257
           :|:||||| | ||  | :  |     ||||  | || |:  |   ||    ||| |||
Db   200  IRKKEDLAREVTEAKREAKAAFGKDEVYLEKLVERARHVESQILGDTHGNVHLFERDCS  259

Qy   258  LQRRHQKVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVD-EKGNHVFIEM  316
           :|||:|||| :|   | || |||  :||:|  | :|  || |||| |  ||||| :|
Db   260  IQRRNQKVERAPAPYLSEAQRQELAAYSLKIAAATNYIGAGTVEYLMDADTGKFYFIEV  319

Qy   317  NPRIQVEHTVEEVTEEVTDVLKAQMRLAAGATL--KELGL-TQDKIKTHGAALQCRITTED  373
           |||||||||| :|||||  |  |    ||||    :|||   :| |  ||  ||| |||||
Db   320  NPRIQVEHTVEVTEVTGIDIVKAQIHILEGAAIGTAESGVPKQEDIRLNGHALQCRITTED  379

Qy   374  PNNGFRPDTGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVAR  432
           |  :  | || || | | | ||| |  ||: |   |    | |  | |       : :|
Db   380  PEHNFIPDYGRITAYRSASGFGIRLDGGTSYTGAVITRYYDPLLVKVTAWAPEPEDAISR  439

Qy   433  AQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILD  492
           | ||:  |   | : |     |  | || ::|  | | | ||  |    | ::    | |
Db   440  MDRALREFRIRGVATNLTFLEAIIGHDSFRNNTYTTRFIDSTPELEFAQVKRQDRATKLLT  499

Qy   493  YLADVTVN---KPHG-VRPKDVAA-PIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLRE  547
            ||| :|    ||   | | |     |  |  |  |       | |  |     |  | |
Db   500  YLADVTVNGHPETKGRAKPADKAAKPI--VPYI-DAPTPDGTKQLLDKLGPKGFADWMRN  556

Qy   548  QDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFL  607
            | | |||||||||||| ||| | |  ::: |   ||| |  |  ||||  ||||  |  |
```

```
Db     5  KKLLVANRGEIAIRIFRAAAELDISTVAIYSNEDKSSLHRYKADESYLVGSDLGPAESYL  64

Qy    90  DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT  149
           :|: ||  |||||||||:||:|||||||||||||||||:||||||||  |: :||    |
Db    65  NIERIIDVAKQANVDAIHPGYGFLSENEQFARRCAEEGIKFIGPHLEHLDMFGDKVKART  124

Qy   150  AAKKAGLPVL-AESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKLA  208
           |  : | ||   :  |  ||  |:|||| ::||||  :|: : || :|:|  |  | | |
Db   125  TAIKADLPVIPGTDGPIKSYELAKEFAEEAGFPLMIKATSGGGGKGMRIVRESELEDAF  184

Qy   209  TEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVHLYERDCSLQRRHQKVEI  268
            |||  | |  | | |    :::|||| |||:||||||| :|  |     ||  ||:|
Db   185  HRAKSEAEKSFGNSEVYIERYIDNPKHIEVQVIGDEHGNIVHLFERDCSVQRRHQKVKVEV  244

Qy   269  APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLV--DEKGNHVFIEMNPRIQVEHTV  326
           |   |  |  |    ||  |    || : | ||       ::| | |||:|||  ||||
Db   245  APSVGLSPTLRQRICDAAIQLMENIKYVNAGTVEFLVSGDE--FFFIEVNPRVQVEHTI  301

Qy   327  TEEVTEVDLVKAQMRLAAGATL--KELGLTQDK-IKTHGAALQCRITTEDPNNGFRPDTG  383
            | :|||||:||: |  |||| :  |:||| ||  |:| ||:| ||||||| |: |  | |
Db   302  TEMVTGIDIVKTQILVAAGADLFGEEINMPQQKDITTLGYAIQCRITTEDPLNDFMPDTG  361

Qy   384  TITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTV  442
            |||||||| || :|    | |  ||   | |:|  | |     :||: |:| | |  :
Db   362  TIAYRSSGGFGVRLDAGDGFQGAEISPYYDSLLVKLSTHAISFKQAEEKMVRSLREMRI  421

Qy   443  SGVAINIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKP  502
            |||  ||||| |  | |::     | | |  | |   :| |  |  :  | |::     |
Db   422  RGVKTNIPFLINVMKNKKFTSGDYTTKFIESTPELFDIQPSLDRGTKTLEYIGNVTIN-  479

Qy   503  HGVRPKDVAAPIDKLPNIKDLPLP---------RGSRDRLKQLGPAAFARD  544
            | |  | ||  |  :| :            |||  | :| :|||||| |
Db   480  ---------GFPNVEKRPKPDYELASIPTVSSSKIASFSGTKQLLDEVGPKGVAEW  526

Qys  545  LREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAM  604
           :::| |  :|||||||||| ||| | |  ::: |   ||| |  |  ||||  ||||
Db   527  VKKQDDVLLDTTFRDAHQSLLATRVRTKDMINIASKTADVFKDGFSLEMWGGATFDVAY  586

Qy   605  RFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKERASSGVDIFRI  664
           | |    |:| |: |  :| |    :||:: : |   : :   || |   | || |:||
Db   587  NFLKENPWERLERLKKAIPNVLFQMLLRASNAVGYKNYPDNVIHKFVQESAKAGIDVFRI  646

Qy   665  FDALNDVSQMRPAIDAVLEINTAVAEVAMAYSGDLSDP-NEKLYTLDYYLKMAEEIVKSG  723
            | :::|   | || ||: |:  | || | |   |:     ||    | :|| |: | |
Db   647  FDSLNWVDQMKVANEAVQEAG-KISEGTICTGDIINPERSNIYTLEYYVKLAKELEREG  705

Qy   724  AHILAIKDMAGLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVD  783
           :|:|:|||||||| |:| | ::    | | |  :|||:| |||| |||  |  ||:| :
Db   706  FHILAIKDMAGLLKPKAAYELIGELKSAVDLPIHLHTHDTSGNGLLTYKQAIDAGVDIID  765

Qy   784  GASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGP  843
           || || || ::|: ||  | |||  |||| ||  :|::|||:||  ||:|:  ||  ||
Db   766  TAVASMSGLTSQPSANSLLYALNGPRHLRTDIEGMESLLSHYWSTVRTYYSDFESDIKSP  825

Qy   844  TGRVYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDL  903
            |||| |:|||| |:|||  | || |:||  ::|::|   : ||||| |||| | ||||
Db   826  NTEIYQHEMPGGQYSNLSQQAKSLGLGERFDEVKDMYRRVNFLFGDIVKVTPSSKVVGDM  885

Qy   904  ALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPGGWPEPLRTRALEGRSEGKAP  963
           ||::       | |    |:  : | :|  ||  || |  |   |  |  | ||  :|
Db   886  ALYMVQNDLDEQSVITDGYKLDFPESVVSFFKGEIGQPVNGFENKDLQAVILKGQE----A  941

Qy   964  LTEVPEEEQAHLDADDSKE-----------RRNSLNRLLRFPKPTEEFLEHRRRFGNTSAL  1012
            |:|     | |  || :|           ||:|:|| : | ||   |  ||:||||:|
Db   942  LTARPGEYLEPVDFEKVRELLEEEQQGPVTEQDIISYVLYPKVYEQYIQTRNQYGNLSLL  1001

Qy   1013 DDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDR  1072
            | || || ||: |:     |   ||| : | :    :|  ::| :  ||  | |  :|
Db   1002 DTPTFFGMRNGETVEIEI-DKGKRLIIKLETISEPDENGNRTIYYAMNGQARRIYIKDE  1060

Qy   1073 SVESVTATAEKADSSNKGHVAAPFAGVVT-VTVAEGDEVKAGDAVAIIEAMKMEATITAS  1131
            :|: |   :||:::||  || :|| | |  ||  :| ||| |  ::|  ||| :|:
Db   1061 NVHTNANVKPKADKSNPSHIGAQMPGSVTEVKVSVGETVKANQPLLITEAMKMETTIQAP  1120
```

```
Db   557 EKRVLVTDTTMRDGHQSLLATRVRTHDIARVASVYSKALPQLLSLECWGGATFDVSMRFL 616
                                         :::::|| ||: |  ||
Qy    80 AKDMGPIESYLSIEEVIRVAKLSGADAIHPGYGLLSESPEFVEACNKAGITFIGPTPDTM 139

Db   608 FEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDA 667
            |||||| :|| :|| |: :|||||| ||  ||||||||||||||||||:|: |||||||
Qy   139 DLTGDKSRAVTAAKKAGLPVLAESTP-SKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRF 197

Db   617 TEDPWERLSLIREGAPNLLLQMLLQMLLRGANGVGYKNYPDNVKYFVRQAARGGVDLFRVFDC 676
            :::::|| ||: |  ||
Qy   140 RQLGNKVAARNLAISVDVPVVPATNLPDDIAEVERMAEEIGYPVMLKASWGGGGRGMRA 199

Db   668 LNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHIL 727
            |: |  |  :|| ||:|||||||||||||||||:|| ||||||||||||||:|:|||||
Qy   198 VSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCS 257

Db   677 LNWVENMRVSMDAIABEN KLCEATICYTGDLLNSARPKYDLKYYTNLAVELEKAGAHI 735
            : :|:|||||||| :|| : |  :|: |
Qy   200 IRKKEDLAREVTEAKREAKAAFGKDEVYLEKILVERARHVESQILGDTHGNVHLFERDCS 259

Db   728 AIKDMAGLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASA 787
            |:||||||||||||||||||:|| ||:||||||||||||||||| ||||||||:|||||
Qy   258 LQRRHQKVELAPAQHLDPELRDRICADAVKFCRSIGYQAGTVEFLVD-EKGNHVFIEM 316

Db   736 AVKDMAGLLKPAAAKVLFKALREATGLPIHFHTHDTSGISAATVLAAVDAGVDAVDAAMD 795
            |:|||||||||||| | |  :|: |||||||| |   |  :|
Qy   260 IQRRNQKVVERAPAPYLSEAQRQELAAYSLKIAAATNYIGAGTVEYLMDADTGKFYFIEV 319

Db   788 PLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRV 847
            |  |   ||:  |  :| : ||:| :: ||||||||
Qy   317 NPRIQVEHTVTEEVTEVDLVKAQMRLAAGATL--KELGL-TQDKIKTHGAALQCRITTED 373

Db   796 AFSGNTSQPCLGSIVEALSGSSERDTGLDTEWIRRISFYWEAVRNQYAAFESDLKGPASEV 855
            |  |   ||:  |   |   |||||:|:|  |: | : | :|  |
Qy   320 NPRIQVEHTVEVVTGIDIVKAQIHILEGAAIGTAESGVPKQEDIRLNGHALQCRITTED 379

Db   848 YRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHL 907
            |||||||| |||||  |  |:|||  ||||:|| |||| || ||:||||||||||| :|
Qy   374 PNNGFRPDTGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVAR 432

Db   856 YLHEMPGGGQFTNLKEQARSLGLESRWHEVAQAYADANRMFGDIVKVTPSSKVVGDMALMM 915
            ||||||||| |||||  |  |||||||||:|: | || |||:|||:||||||||||:||:|
Qy   380 PEHNFIPDYGRITAYRSASGFGIRLDGGTSYTGAVITRYYDPLLVKVTAWAPEPDEAISR 439

Qy   908 VGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEV 967
            |  | || ||||| |||:|||| :||: |: |||||||||:| |||:|||:|||:|||||
Db   916 VSQDLTVADVENPDREVSFPDSVVSMLKGDLGQSPGGWPEALQKKAL---KGEKPYTVR 971

Qy   968 PEE--EQAHLDADDSKERRNSLNR------LLFPKPTEEFLEHRRRFGNTSALDDR 1015
            |||   || || |:|:|| | ||       | || |::| ||:|||:|::|::|||
Db   972 PGSLLEDADLDA-ERKVIETKLERKVDDFEFASYLMYPKVFTDFALTAETYGPVSVLPTH 1030

Qy  1016 EFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRS-V 1074
            |||:|| |::| || ||||:|:||||  ||| ||:||:|:| |::|  || |:|||:||
Db  1031 AYFYGMEDGEELFADIERGKTLVIVN-QASSGIDDKGMVTVFFEINGQPRRIKVPDRAHG 1089

Qy  1075 ESVTATAEKADSSNKGHVAAPFAGVVT-VTVAEGDEVKAGDAVAIIEAMKMEATITASVD 1133
            ||||| |::||:| ||:|| |  ||| ||| ||||||||||||||:|||||||:| |  |
Db  1090 ASGSAVRRKAEPGNASHIGAPMPGVISRVFINGQGEVKAGDVLLSIEAMKMETALHAERD 1149

Qy  1134 GKIERVVPAATKVEGGDLIV 1154
            ||:| |  |   |:|||:||
Db  1150 GKIAEVLVKPGDIDAKDLLI 1170
```

RESULT 12

C97686

pyruvate carboxylase (U51439) [imported] - Agrobacterium tumefaciens (strain C58, Cereon
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C;Accession: C97686
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: C97686
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1174 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK88444.1; PID:g15157941; GSPDB:GN00169
C;Genetics:
A;Gene: AGR_C_4940
A;Map position: circular chromosome
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

Query Match   42.2%; Score 2474.5; DB 2; Length 1174;
Best Local Similarity   46.5%; Pred. No. 3.9e-128;
Matches 540;   Conservative 174;   Mismatches 406;   Indels 41;   Gaps 19;

```
Qy    25 TLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFPASEAVRIG----- 79
            ||   |||||||||||:|| |||||:|| ||:|| |: ||: :|: |  |:|
Db    20 TVLKISKILVANRSEIAIRVFRAANELGIKTVAIWAEEDKLSLHRFKADESYQVGRGPHL 79

Qy    80 -TEGSPVKAYLDIEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPIPEVL 138
              |||| | ||| ::|  |  | |  |||| || : |  || |  |||||||| ||||
```

Db   1150 GKIAEVLVKPGDQIDAKDLLI 1170

RESULT 13

AE3285

Pyruvate carboxylase (EC 6.4.1.1) [imported] - Brucella melitensis
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C;Accession: AE3285
R;Delvecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AE3285
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1158 <KUR>
A;Cross-references: GB:AE008917; PIDN:AAL51448.1; PID:g17982157; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI0266
A;Map position: I
C;Superfamily: Pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding
C;Keywords: ligase

Query Match   42.0%;   Score 2465.5;   DB 2;   Length 1158;
Best Local Similarity   46.5%;   Pred. No. 1.2e-127;
Matches   537;   Conservative   172;   Mismatches   408;   Indels   37;   Gaps   15;

Qy   31 KILVANRGEIAVRAFRAALETGAAVAIYPREDRGSFHRSFASEAVRIG-----TEGSP 84
       |||| ||| ||||||| : |: ||| || | :| :: || ||
Db   10 KILVANRSEIAIRVFRAANELGLKTVTIWAEEDKLSLHRFKADESYQVGRGPHLDRDLGP 69

Qy   85 VKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDK 144
       :|||| |||| | ||| |||||||||||| : |||:|| ||||||||| ||||:||||
Db   70 IESYLSIDEIIRVAKLSGADAIHPGYGLLSESPEFAEACAENGIVFIGPKPETMRRLGNK 129

Qy   145 SRAVTAAKKAGLPVLAESTP-SKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDE 203
       : : |: : | :::::: | / :|: ||| ||||||||||||: : :
Db   130 VAARNLAIEIGVPVVPATDLPLPDDMDEVKKLAAQIGYPLMLKASWNGGGGRGMRAIRAEAD 189

Qy   204 LRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQ 263
       |||||||||||||||||| :||||||| |::||||||||| :|:|||:|
Db   190 LAREVMEAKRAKAAFGKDEVYLEKLVERARHVEVQILGDTYGNAVHLFERDCSIQRRNQ 249

Qy   264 KVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVD-EKGNHVFIEMNPRIQV 322
       ||| |||| | | ||| |||||||| || |||||||| ||: ||||||||||||
Db   250 KVVERAPAPYLNDAQRRELADYGLKIAHATDYIGAGTVEFLMDADTGKFYFIEVNPRIQV 309

Qy   323 EHTVTEEVTEVDLVKAQMRLAAGATL--KELGL-TQDKIKTHGRALQCRITTEDPNNGFR 379
       |||||||| :|:||||| :|:: :| | | :|: ||||||||||||||||
Db   310 EHTVTEEVTGIDIVKAQIHILEGFAIGTPESGVPRQEDIRLNGHALQCRITTEDPEQNFI 369

Qy   380 PDTGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALA 438
       |||||||||||||||| ||||||||| |||||||
Db   370 PDYGRIQAYRSAAGFGIRLDGGTAYSGAFITRYYDIARIATGFIGDHPHLLQAPPADDEQGRILDYLADVT 429

Qy   439 EFTVSGVAINIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVT 498
       |:: ||||| :|||| : |: :| : |||||||||||||||||||||||||
Db   430 EFRIRGVAINLTFLEAIINHPKFLSNDYTTRFIDTTPELFEQMRKQDRATKLLIYIADVT 489

Qy   499 VN-----KPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAV 553
       || | :: | ||||||||| | |:|||
Db   490 VNGHPETKGRAKPARDAAKP--RVPWFGDKIVADGTKQLLDLQLGPKKFAEWMRNEKRALI 547

Qy   554 TDTIFPRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWD 613
       | |:|||||||||||||||: |||||| :| |: :|||| | ||||
Db   548 TDTTMRDGHQSLLATRVRTYDIARIANAYAQALPNLFSLECWGGATFDVSMRFLTEDPWE 607

Qy   614 RLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQ 673
       || |: |||||:| | ||||||| | || ||| || |||||||
Db   608 RLALVREGAPNLLQMLLRGANGVGYKSYPDNVVKYFVREAARAGIDLFRVFDSLNWVEN 667

Db   674 MRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYILKMAEEIVKSGAHILAIKDMA 733
       || :||||| ||| : :|: :|| ::|: :|::: ||| |||
Qy   668 MRVSMDAVLEEN-KLCEAAICYTGDILNPDRAKYDLNYYVNLAKEVEKAGAHIIAVKDMA 726

Qy   734 GLLRPAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTT 793
       |||||||| ||: | |||:||||:| ||||||: ||| |: |||||
Db   727 GLLKPAAARVLFKALREETDLPIHFHTHDTSGISAATVLAAIDAGVVDAAMDALSGNT 786

Qy   794 SQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIP 853
       ||| || |: ::|:| | ||: || :||||| :|: : |||||
Db   787 SQPCLGSIVEALNGSERDSGLDPDLIRRISFYWEAVRHQYAAFESDLKGPASEVYLHEMP 846

Qy   854 GGGLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPIKVTPSSKVVGDLALHLVGAGVD 913
       | | : ||| || :||||||||| ||||||||||| : |
Db   847 GGQFTNLKEQARSLGLETRWHEVAQAYADVNRMFGDIVKVTPSSKVVGDMALMVAQDLT 906

Qy   914 PADFAADPQKYDIPDSVIAFLRGELGNPPGGWEPLRTRALEGRSEGKAPLTEVPEE--E 971
       ::|| : : || |:
Db   907 VADVENPAKDIAFPDSVVSMRGDLGQPPSGWPEALQKKVL----KDEKPFTVRPGSLLP 962

Qy   972 QAHLDA------DDSKERRNS---LNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLV 1022
       |||| :: | || :::| :|:|
Db   963 AADLLDAERKSFEDSVGRKLSDQEFASALMYPKVFTDYATAHETYGPTSVLPTPVYFYGLK 1022

Qy   1023 EGRETLIRLPDVRTPLLYVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAE 1082
       :|: : | :|:
Db   1023 PEEEVFVDLERGKTLVIVN-QAMSETDEKGMVTVFFELNGQPPRIKVPNRAKGASGGVRR 1081

Qy   1083 KADSSNKGHVAAPFAGVV-TVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVV 1141
       | ||| | : || |: | ::| ||||
Db   1082 KVEAGNDKQVGAPMPGVISTVAVVAGQKVTQGDVLLSIEAMKMETAIHAERDGTIAEVLV 1141

Qy   1142 PAATKVEGGDLIVV 1155
       :: ||||| |:
Db   1142 RPGEQIDAKDLLIV 1155

RESULT 14

T20346

pyruvate carboxylase (EC 6.4.1.1) D2023.2 [similarity] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000
C;Accession: T20346
R;Kershaw, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19260
A;Accession: T20346
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1175 <WIL>
A;Cross-references: EMBL:Z81052; PIDN:CAB02872.1; GSPDB:GN00023; CESP:D2023.2
A;Experimental source: clone D2023
C;Genetics:
A;Gene: CESP:D2023.2
A;Map position: 5
A;Introns: 34/3; 103/3; 246/1; 451/3; 567/2; 821/1; 1045/3; 1126/3
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindi
C;Keywords: ligase
F;1140/Binding site: biotin (Lys) (covalent) #status predicted

Query Match   42.0%;   Score 2461;   DB 2;   Length 1175;
Best Local Similarity   45.1%;   Pred. No. 2.2e-127;
Matches   522;   Conservative   201;   Mismatches   389;   Indels   46;   Gaps   14;

Qy   29 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 88
       | :|:||||||||||| : :| || : | |||||
Db   32 FNKVMVANRGEIAIRVFRALTELNKTSVAIYAEQDKNSMHRLKADEAYLVGKGLPVVAY 91

Qy   89 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 148
       ||| ||||:|||||||| |||||||||||| ||| |||
Db   92 LTIDQIIETALKHNIDAIHPGYGFLSERSDFAAACQNAGIVFIGPSPDVMARMGDKVAAR 151

```
RESULT 15
QYBYP

pyruvate carboxylase (EC 6.4.1.1) 1 [validated] - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein G3428; protein YGL062w: pyruvic carboxylase
C;Species: Saccharomyces cerevisiae
C;Date: 31-Dec-1991 #sequence_revision 19-Jul-1996 #text_change 01-Feb-2002
C;Accession: S64066; A29233; S05760; A29722
R;Feuermann, M.; Potier, S.; Souciet, J.L.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64044
A;Accession: S64066
A;Molecule type: DNA
A;Residues: 1-1178 <FEU>
A;Cross-references: EMBL:Z72584; NID:g1322565; PIDN:CAA96765.1; PID:g1322566; GSPDB:G
A;Experimental source: strain S288C
R;Lim, F.; Morris, C.P.; Occhiodoro, F.; Wallace, J.C.
J. Biol. Chem. 263, 11493-11497, 1988
A;Title: Sequence and domain structure of yeast pyruvate carboxylase.
A;Reference number: A92662; MUID:88298805; PMID:3042770
A;Accession: A29233
A;Molecule type: DNA
A;Residues: 1-461,'G',463-492,'D',494-594,'A',596-618,'Q',620-663,'S',665-771,'R',773
A;Cross-references: EMBL:J03889; NID:g172101; PIDN:AAA34843.1; PID:g172102
R;Morris, C.P.; Lim, F.; Wallace, J.C.
Biochem. Biophys. Res. Commun. 145, 390-396, 1987
A;Title: Yeast pyruvate carboxylase: gene isolation.
A;Reference number: S05760; MUID:87241529; PMID:3036126
A;Accession: S05760
A;Molecule type: DNA
A;Residues: 1003-1178 <MOR>
A;Cross-references: EMBL:J03889
A;Accession: A29722
A;Molecule type: protein
A;Residues: 1124-1149 <MOR2>
C;Genetics:
A;Gene: SGD:PYC1; PYV; MIPS:YGL062w
A;Cross-references: SGD:S0003030; MIPS:YGL062w
A;Map position: 7L
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindi
C;Keywords: biotin binding; gluconeogenesis; homotetramer; ligase; zinc
F;21-478/Domain: biotin carboxylase homology <BCH>
F;157-331/Domain: ATP/bicarbonate binding #status predicted <ATB1>
F;353-468/Domain: ATP/bicarbonate binding #status predicted <ATB2>
F;569-908/Domain: pyruvate binding #status predicted <PYR>
F;1096-1169/Domain: lipoyl/biotin-binding homology <LPB>
F;1135/Binding site: biotin (Lys) (covalent) #status experimental

Query Match            41.9%;  Score 2457;  DB 1;   Length 1178;
Best Local Similarity  46.6%;  Pred. No. 3.6e-127;
Matches  540;  Conservative 175;  Mismatches 400;  Indels 44;  Gaps 16;

Qy    31 KILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEG--SPVKAY 88
          ||||||||||  || : ||   : ||||| ||: || :::::|||: ||| :   :|||
Db    21 KILVANRGEIPIRFRTAHELSMQTVAIYSHEDRLSTHKQKADEAVVIGEVGQYTPVGAY 80

Qy    89 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 148
          | |||| | |: |:|  |:||||||||| : |   : ||||:||  |||:|: ||||: |
Db    81 LAIDEIISIAQKHQVDFIHPGYGFLSENSEFADKVVKAGITWIGPPAEVIDSVGDKVSAR 140

Qy   149 TAAKKAGLPVLAEST-PSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKL 207
          : |||| |||       : ||:|: | ||  |  ||:||  |||||||||||:   :|:|
Db   141 NLAAKANVPVPGTPGPIETVEEALDFVNEYGYPVIKAAFGGGGRGMRVVREGDDVADA 200

Qy   208 ATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVE 267
          :   :  : ||:||   :||| ||  |||||| | |   | ||| ||||::|||||||||
Db   201 FQRAITSEARTAFGNGTCFVERFLDKPKHIEVQLLADNHGNVVHLFERDCSVQRRHQKVVE 260

Qy   268 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 327
          :||| :| ||::|  | ||||  :  ||: ||| |||||:: | :||| |||||||||:|
Db   261 VAPAKTLPREVRDAILTDAVKLAKECGYRNAGTAEFLVDNQNRHYFIEINPRIQVEHTIT 320

Qy   328 EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 387
```

```
Qy   149 TAAKKAGLPVLAEST-PSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKL 207
          : |||| |||  :    : :|::|:   |: |  ||:||  ||||||::|:| |:::  :
Db   152 QAAIEAGVQVVPGTPGPITTADEAVEFAKQYGTPIILKAAYGGGGRGIRRVDKLEVEEA 211

Qy   208 ATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVE 267
          :   | :   ||||||  |||:| : | |||| |||   |:|||||||||::|||||||||
Db   212 FRRSYSEAQAQAAFGDGSLFVEKFVERPRHIEVQLLGDHHGNIVHLYERDCSVQRRHQKVVE 271

Qy   268 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 327
          :||| :| || | | : ||::  :  ||:| ||||||||:: | ::|| | |:||||||:|
Db   272 LAPAPALPEGVREKILADALRLARHVGYQNAGTVEFLVDQKGNYFIEVNARLQVEHTVT 331

Qy   328 EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 387
          ||:| |||| | ||::  ::|::|  :| :||:| | :||| ||||| ||| |||:| |
Db   332 EEITGVDLVQAQIRITAEGKSLDDLKLSQETIQTTGSAIQCRVTTEDPAKGFQPDSGRIEV 391

Qy   388 YRSPGGAGVRLDGAAQLGGE-ITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVA 446
          :|  |  |:||| :|  :|  ::: :||:|||  :   : | || ||||::::||:: ||
Db   392 FRSGEGMGIRLDSASAFAGSVISPHYDSLMVKVIASARNHPNAAAKMIRALKKFRIRGVK 451

Qy   447 TNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKP---- 502
          |||  | :|||: ||  :  | :|||:||:|:|      ::|::|:||:::|:|
Db   452 TNIPFLLNVLRQPSFLDASVDTYFIDEHPELFQFKFPSQNRAQKLLNYLGEVKVNGPTTPL 511

Qy   503 -HGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDA 561
           : | | || :|   | :|    |:|   |:|   ||:  |::  |    :||||||||
Db   512 ATDLKPAVVSPPIPYIP-AGAKPPTGLRDVLVQRGPTEFAKEVRSRPGCMITDTTFRDA 569

Qy   562 HQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREA 621
          |||||||||| :   :    ::   |:|:||| ||||::|::|||| | ||:||  :|:|
Db   570 HQSLLATRVRTYDMAAISPFVAQSFNGLFSLENWGGATFDVSMRFLHECPWERLQTLRKL 629

Qy   622 MPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAV 681
          :|:: :|:|||| | :|| ::|| | : ||:  |: | |:||:||:||:: ::   :: |
Db   630 IPNIPFQCLLRGANAMGYSNYPDNVIYKFCELAVKNGMDVFRVFDSLNVLPNLLVGMEAV 689

Qy   682 LETNTAVAEVAMAYSGDLSDPNEKLYTLDYLKMAEEIVKSGAHILAIKDMAGLLRPAAV 741
           :   : :| ||| |||::|   : |:|:||::|::::||:|||:|:||||| |:| | |
Db   690 GKAG-GVVEAAIAYTGDVTDKSRDKYDLKYYLNLADQLVKAQAHILSIKDMAGVLKPEAA 748

Qy   742 TKLVTALRREF-DLPVHVHTHDTRAGGQLATYFARAAQAGADAVDGASAPLSGTTSQPSLSA 800
          .|| .||: :| |:|:|||||||: :|:|| ::  | :||||:||  : |:||||||||: |
Db   749 KLLIGALRDKFPDIPIHVHTHDTSGAGVAAMLECAKAGADVVDAAVDSMSGMTSQPSMGA 808

Qy   801 IVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGIYLPFESGTPGPTGR--VYRHEIPGGQLS 858
          |||::: |::|||||| :: ::| |||:: :|||||| | :  ||:  |||||||||| :
Db   809 IVASLQGTKHDTGLSLDDISKYSAYWESTRQLYAPFECATTMKSGNADVYKHEIPGGQYT 868

Qy   859 NLRAQATAIGLGADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHIVGAG-----V 912
          || .|| ::|||: :|:::| ||  :||:| | ||:||||:|||||| |:|  :     |
Db   869 NLQFQAFSLGLGPQFDEVKRMYREANLVLGDIIKVTPSSKIVGDLAQFMVQNNLTRETLV 928

Qy   913 DPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRS-EGKAPLTEVPPE 970
          | ||        :||| |: |::| :|:|| |:|||: :  | |:    ||  :: |
Db   929 DRAD------DLSFPKSVVDFMQGNVGQPPYGFPEPLRIKVLRGKPKVDGR------PGE 976

Qy   971 EQAHLDADDSK--------ERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFY 1019
          | ::|| :  :         ::: |  ::  ::|:|::      |: ::  :  :
Db   977 NAKPVDLDAVKVELEEKHGRTLSEDVMSYSMFFTVFDFEETFRQQYGPVDKLFTRLFLT 1036

Qy  1020 GLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTA 1079
          || :  | : ::   ::|   :||: :    ::| :::| : :::| :::   ::  |:
Db  1037 GLEIAEEVDVEIRSGKT-LAIQLLAEGKLNKRGEREVFDLNGQMRSIFVVDKEASKEIV 1095

Qy  1080 TAEKADSSNKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIER 1138
          | ::      ||: || || |: : :: || |: :: | : :||||| :| :|:  |:
Db  1096 TRPRALPGVRGHIGAPMPGDVLEIKIKEGDKVIKKQPLFVLSAMKMEMVIDSPIAGTVKA 1155

Qy  1139 VVVPAATKVEGGDLIVVV 1156
           :  ::|: :|||::||:
Db  1156 IHAPQGTKCSAGDLVVEV 1173
```

```
Db   321  EEITGIDIVAAQIQIAAGASLPQLGLFQDKITTRGFAIQCRITTEDPAKNFQPDTGRIEV 380

Qy   388  YRSPPGGAGVRLDGA-AQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVA 446
          || ||||||| :  ||:  ||::|| ::|::|::|    || || || ||||::| ||
Db   381  YRSAGGNGVRLDGGNAYAGTIISPHYDSMLVKCSCSGSTYEIVRRKMIRALIEFRIRGVK 440

Qy   447  TNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVN----KP 502
          |||  |: || |    |||   |:   || |:: |||| |  ||||||||||| |   ||
Db   441  TNIPFLLTLLTNPVFIEGTYWTTFIDDTPQLFQMVSSQNRAQKLLHYLADVAVNGSSIKG 500

Qy   503  HGVRPKDVAAPIDKLPNIKDL---------PLPRGSRDRLKQLGPAAFARDLREQDALA 552
          ||  |  |:   :|| || |           |  |  || :|||| | : |: :: |:
Db   501  QIGLPKLKSNP--SVPHLHDAQGNVINVTKSAPPSGWRQVLLEKGPAEFARQVRQFNGTL 558

Qy   553  VTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPW 612
          ::||||||||||||||||| ||| :|| ||| |  :|||:||| |||:|||||||:|||||
Db   559  LMDTTWRDAHQSLLATRVRTHDLATIAPTTAHALAGRFALECWGGATFDVAMRFLHEDPW 618

Qy   613  DRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVS 672
          :|: ||  |||| |||||||||  | | :|: | |  ||| |||:||||||||||||||:||
Db   619  ERLRKLRSLVPNIPFQMLLRGANGVAYSSLPDNAIDHFVKQAKDNGVDIPRVFDALNDLE 678

Qy   673  QMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYIKMAEEIVKSGAHILAIKDM 732
          |:|   |||:|  | :|||:  ::|| | |||| :|| :|| ||| ||::|||||||||||||
Db   679  QLKVGVDAVKKAG-GVVEATVCFSGDMLQPGKK-YNLDYYLEIAEKIVQMGTHILGIKDM 736

Qy   733  AGLLRPAAVTKLIVTALRREF-DLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSG 791
          |||:: ||  |||| |  ||  :|: || |||||||||:|  ||::||||||||| |  |
Db   737  AGTMKPAAAKLLIGSLRAKYPDLPIHVHTHDSAGTAVASMTACALAGADVVDVAINSMSG 796

Qy   792  TTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHE 851
          |||||||| ::::|||  | | || | :  | |::| :  ||: ||:  |  |  || :|
Db   797  LTSQPSINALLASL-EGNIDTGINVEHVRELDAYWAEMRLLYSCFEADLKGPDPEVYQHE 855

Qy   852  IPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAG 911
          |||||| |||:|| |  |||| |:| ::: || :| || |:||||||||||||||||||||:|
Db   856  IPGGQLTNLLFQAQQLGLGEQWAETKRAYREANYLGDIVKVTPTSKVVGDLAQFMVSNK 915

Qy   912  VDPADFAADPQKTDIPDSVIAFLRGELGNPPGGWNPEPLRTRALEGRSEGKAPLTEVPEEE 971
          ||| :|| || || ||| |::  || ||||||||| ||  |  :|| |
Db   916  LTSDDVRRLANSLDFPDSVMDFFEGLIGQPYGGFPEPFRSDVLRNK--RRKLTCRPGLE 972

Qy   972  QAHIDADDSKERRNSLNRL---------LFPKPTEEFLEHRRRFGNTSALDDREFFY 1019
          :|: :|  |:  |: ||  :    : :|  :  |  ||||||||| : ::  |: :
Db   973  LEPPDLE--KIREDLQNRFGDVDECDVASYNMYPRVYEDFQKMRETYGDLSVLPTRSFLS 1030

Qy   1020 GLVEGRETLIRLPDVRTPLLVRLDAISEPDDK-GMRNVVANVNGQIRPMRVRDRSVESVT 1078
          ||: :| |: | |  |  |:  |::   || ||:| |:  |:||  |:||:| ||:|  |
Db   1031 PLETDEEIEVVIEQGKT-LIIKLQAVGDLNKKTGEREVYFDLNGEMRKIRVADRSQKVET 1089

Qy   1079 ATAEKADSSNKGHVAAPFAGV-VTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIE 1137
          ||  | ||:| ||  |  | |   :|  :| :|| |:  ||:|||||||||:|| |||||:
Db   1090 VTKSKADMHDPLHIGAPMAGVIVEVKVHKGSLIKKGQPVAVLSAMKMEMIISSPSDGQVK 1149

Qy   1138 RVVVPAATKVEGGDLIVVV 1156
          ||::|  |:|| ||:|:|:
Db   1150 EVFVSDGENVDSSDLIVLL 1168
```

Search completed: September 24, 2003, 16:00:27
Job time : 40.259 secs

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:   September 24, 2003, 15:44:03 ; Search time 18.6369 Seconds
                                        (without alignments)
                            2919.469 Million cell updates/sec

Title:    US-09-974-973A-2
Perfect score: 5865
Sequence: 1 MTAITLGGLLLKGIIILVST.........RVVVPAATKVEGGDLIVVVS 1157

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:  127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters:     127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SwissProt_41:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description	
1	2550.5	43.5	1178	1	PYC_MOUSE	Q05920	mus musculu
2	2534.5	43.2	1178	1	PYC_RAT	P52873	rattus norv
3	2531.5	43.2	1178	1	PYC_HUMAN	P11498	homo sapien
4	2457	41.9	1178	1	PYC1_YEAST	P11154	saccharomyc
5	2436	41.5	1180	1	PYC2_YEAST	P32327	saccharomyc
6	2393	40.8	1189	1	PYC_PICPA	P78992	pichia past
7	1079	18.4	501	1	PYCA_METJA	Q58626	methanococc
8	1032.5	17.6	447	1	ACCC_ANASP	Q06862	anabaena sp
9	1020	17.4	506	1	PYCA_ARCFU	O30019	archaeoglob
10	996	17.0	491	1	PYCA_METTH	O27939	methanobact
11	950	16.2	448	1	ACCC_HAEIN	P43873	haemophilus
12	940	16.0	449	1	ACCC_ECOLI	P24182	escherichia
13	938	16.0	449	1	ACCC_ECO57	Q8X9P6	escherichia
14	921	15.7	449	1	ACCC_PSEAE	P37798	pseudomonas
15	897	15.3	725	1	MCCA_HUMAN	Q96RQ3	homo sapien
16	894	15.2	717	1	MCCA_MOUSE	Q99MR8	mus musculu
17	891.5	15.2	654	1	BCCA_MYCTU	P46401	mycobacteri
18	875	14.9	448	1	ACCC_BACSU	P49787	bacillus su
19	874.5	14.9	567	1	PYCB_METJA	Q58628	methanococc
20	869.5	14.8	598	1	BCCA_MYCLE	P46392	mycobacteri
21	869	14.8	703	1	PCCA_HUMAN	P05165	homo sapien
22	865	14.7	731	1	MCCA_SOYBN	Q42777	glycine max
23	862	14.7	734	1	MCCA_ARATH	Q42523	arabidopsis
24	842.5	14.4	704	1	PCCA_RAT	P14882	rattus norv
25	837.5	14.3	1835	1	DUR1_YEAST	P32528	saccharomyc
26	828	14.1	568	1	PYCB_METTH	O27179	methanobact
27	806	13.7	590	1	DCOA_SALTY	Q03030	salmonella
28	801.5	13.7	595	1	DCOA_KLEPN	P13187	klebsiella
29	620.5	10.6	2345	1	COA1_RAT	P11497	rattus norv
30	617.5	10.5	2324	1	COA1_CHICK	P11029	gallus gall
31	617.5	10.5	2346	1	COA1_HUMAN	Q13085	homo sapien
32	616.5	10.5	2346	1	COA1_BOVIN	Q9TTS3	bos taurus
33	607.5	10.4	2346	1	COA1_SHEEP	Q28559	ovis aries
34	570.5	9.7	2273	1	HFA1_YEAST	P32874	saccharomyc
35	560.5	9.6	2483	1	COA2_HUMAN	O00763	homo sapien
36	547.5	9.3	2233	1	COAC_YEAST	Q00955	saccharomyc
37	519	8.8	2280	1	COAC_SCHPO	P78820	schizosacch
38	267.5	4.6	1076	1	CARB_ARCFU	O28994	archaeoglob
39	266	4.5	1060	1	CARB_METTH	O27077	methanobact
40	255.5	4.4	1077	1	CARB_VIBVU	Q8DEM2	vibrio vuln
41	255	4.4	1072	1	CARB_ECO57	Q8XA38	escherichia
42	254.5	4.3	1099	1	CARB_THEMA	Q9WZ27	thermotoga
43	252.5	4.3	1077	1	CARB_VIBPA	Q87SF3	vibrio para
44	252	4.3	1072	1	CARB_ECOLI	P00968	escherichia
45	248.5	4.2	1074	1	CARB_SALTY	P14846	salmonella

```
                              ALIGNMENTS

RESULT 1
PYC_MOUSE
ID   PYC_MOUSE     STANDARD;      PRT;   1178 AA.
AC   Q05920;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic
DE   carboxylase) (PCB).
GN   PC OR PCX.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Adipocyte;
RX   MEDLINE=93189578; PubMed=8446598;
RA   Zhang J., Xia W.L., Brew K., Ahmad F.;
RT   "Adipose pyruvate carboxylase: amino acid sequence and domain
RT   structure deduced from cDNA sequencing.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:1766-1770(1993).
CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
CC       CARBOXYL GROUP TO PYRUVATE IN THE SECOND. CATALYZES IN A TISSUE
CC       SPECIFIC MANNER, THE INITIAL REACTIONS OF GLUCOSE (LIVER, KIDNEY)
CC       AND LIPID (ADIPOSE TISSUE, LIVER, BRAIN) SYNTHESIS FROM PYRUVATE.
CC   -!- CATALYTIC ACTIVITY: ATP + Pyruvate + HCO(3)(-) = ADP + phosphate +
CC       oxaloacetate.
CC   -!- COFACTOR: BIOTIN AND MANGANESE.
CC   -!- PATHWAY: GLUCONEOGENESIS AND LIPOGENESIS.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC   -!- TISSUE SPECIFICITY: LIVER, KIDNEY, ADIPOSE TISSUE, LIVER
CC       AND BRAIN.
CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC       AND CARBAMYL PHOSPHATE SYNTHETASES.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; L09192; AAA39737.1; -.
DR   PIR; A47255; A47255.
DR   HSSP; P24182; 1BNC.
DR   SWISS-2DPAGE; Q05920; MOUSE.
DR   MGD; MGI:97520; Pcx.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
```

```
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
KW   Ligase; Multifunctional enzyme; Biotin; Manganese; Gluconeogenesis;
KW   ATP-binding; Mitochondrion; Lipid synthesis; Transit peptide.
FT   TRANSIT       1     20       MITOCHONDRION (POTENTIAL).
FT   CHAIN        21   1178       PYRUVATE CARBOXYLASE.
FT   DOMAIN       21    549       BIOTIN CARBOXYLASE (BY SIMILARITY).
FT   DOMAIN      550   1000       CARBOXYLTRANSFERASE (BY SIMILARITY).
FT   DOMAIN     1096   1178       BIOTIN CARBOXYL CARRIER PROTEIN
FT                                (BY SIMILARITY).
FT   NP_BIND     198    203       ATP (BY SIMILARITY).
FT   ACT_SITE    328    328       BY SIMILARITY.
FT   BINDING    1144   1144       BIOTIN (BY SIMILARITY).
SQ   SEQUENCE   1178 AA;  129684 MW;  14CEA0F9DABB8127 CRC64;
```

```
Query Match             43.5%;  Score 2550.5;  DB 1;  Length 1178;
Best Local Similarity   46.1%;  Pred. No. 9e-127;
Matches  543;  Conservative  192;  Mismatches  407;  Indels  35;  Gaps  13;
```

```
Qy      7 GGLLIKGIITLVSTHTSSTLPA----------FKKILVANRGEIAVRAFRAALETGAA 54
            :|:: |                       |:||||||||| ||| || |
Db      9 GGLRLLGV------RRSSAPVASPNVRRLEYKPIKKVMVANRGEIAIRVFRACTELGIR 62

Qy     55 TVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLDIEIIGAAKKVKADAIYPGYGFLS 114
           :|  |   ||  ||||||||| :::||||||||||| || : |::| |
Db     63 TVAVYSEQDTGQMHRQKADEAYLIGRGLAPVQAYLHIPDIIKVAKENGVDAVHPGYGFLS 122

Qy    115 ENAQLARECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVL-AESTPSKNIDDIVK 173
           || :| |: |||: ||||||||  |||      ||| ::|||   ||:
Db    123 ERADFAQACQDAGVRFIGPSPEVVRKMGDKVEARAIAIAAGVPVVPGTDSPISSLHEAHE 182

Qy    174 SAEGQTVYPIFVKAVAGGGRGMRFVSSPDELRKLATEASREAEAAFGDGSVYVERAVINP 233
           ||:| | | || :|: |||||||  ||| | | | : |::|| | :| :| | : |
Db    183 FSNTFGFPIIFKAAYGGGGRGMRVVHSYEELEENYTRAYSEALAAFGNGALFVEKFIEKP 242

Qy    234 QHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSI 293
           :|| ||:||| |  || |||||| |||:|:|||| |||:  |: |:||:|| |::
Db    243 RHIEVQILGDQYGNILHLYERDCSIQRRHQKVVEIAPATHLDPQLRSRLLTSDSVKLIAKQV 302

Qy    294 GYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLIVKAQMRLAAGATLKELGL 353
           || ||||||||||:|||||  |||||||| ||||| |:|| |||| :::: :|: ||:|
Db    303 GYENAGTVEFLVDKHGKHYFIEVNSRLQVEHTVTEEITDVDLVHAQIHVSEGRSLPDLGL 362

Qy    354 TQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQL-GGEITAHF 412
           :| |:    |  |  || |||  | | || | |  | ||: | | |::: | | : |
Db    363 RQENIRINGCAIQCRVTTEDPARSFQPDTGRIEVFRSGEGMGIRLDNASAFQGAVISPHY 422

Qy    413 DSMLVKMICRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIG 472
           |::: | |:  | : ||| || |||||:  || ||||| || || ::| :|| |:||:|
Db    423 DSLLVKVIAHGKDHPTAATKMSRALAEFRVRGVKTNIPFLQNVLNNQQFLAGTVDTQFID 482

Qy    473 DHPHLLQAPPADDEQGRIILDYLADVTVNKPHGVRPKDVA-APID-KLPNIKDLPLPRGSR 530
           :| || |: : :  :   || |  | | :: : :: ::   | :
Db    483 ENPELFQLRPAQNRAQKLLHYLGHVMVNGPTTPIPVNVSPSPVDPAVPVVPIGPPAAGFR 542

Qy    531 DRLKQLGPAAFARDLREQDALAVDTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELL 590
           |: |:|| | :|||: ||||| | | |||||||||||| ||  | || | ||
Db    543 DILLREGPEGFARAVRNHQGLLLMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFNKLF 602

Qy    591 SVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVIQMLLRGRNTVGYTPYPDSVCRAF 650
           | | || |: | | |||  ||  || :||  |||: ||||| | | || |:  |
Db    603 SMENWGGATFDVAMRFLYECPWRRLQELRELIPNIPFQMLLRGANAVGYTNYPDNVVFKF 662
```

```
Qy    651 VKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLD 710
           :  |: ::||::| :||:| :| |       :  : ::|| :|
Db    663 CEVAKENGMDVFRVFDSLNYLPNMLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLE 721

Qy    711 YYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVTHDTAGGQLA 769
           ||||||||::|| | ||||| ||||| | | :| |:| ||| |: || |||:|::
Db    722 YYMGLAEELVRAGTHILCIKDMAGLLKPAACTMLVSSLRDRFPDLPLHIHTHDTSGAGVA 781

Qy    770 TYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAV 829
           | | ||||||||||  ||||| ||  || | ||  |: |:| || ::| ||:|:|
Db    782 AMLACAQAGADVVDVADSMSGMTSQPSMGALVACTKGTPLDTEVPLERVFDYSEYWEGA 841

Qy    830 RGLYLPFESGTPGPTGR--VYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEML 887
           ||| | | :    :|      : :: |  :: |:: |  | | ::| : |||| ||
Db    842 RGLYAAFDCTATMKSGNSDVYENEIPGGQYTNLHFQAHSMGLGSKFKEVKKAYVEANQML 901

Qy    888 GRPTKVTPSSKVVGDLALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPE 947
           | | |||||| ||:| |  || || :  |   || || ||||  |:|| ||:|| ||  |
Db    902 GDLIKVTPSSKIVGDLAQFMVQNGLSRAEAEAQAEELSFPRSVVEFLQYGIGIPHGGFPE 961

Qy    948 PLRTRAL-----EGRSEGKAPLTEVPEEQAHLDA-DDSKERRSLNRLLFPKPTEEFL 1000
           || :|:      ||: ||  || |  :|:
Db    962 PFRSKVLKDLPRIEGRPGASLPPINLKELEKDLIDRHGEEVTPEDVLSAAMYPDVFAQFK 1021

Qy   1001 EHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANV 1060
            :: :  |: ::|:|  |   ::|:| ||: :|:| ::| | | : | |  |:|| :|
Db   1022 DFTATFGPLDSLNTRLFLQGPKIAEEFVELERGKT-LHIKALAVSDLNRAGQRQVFFEL 1080

Qy   1061 NGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAII 1119
           ||| |: : :| |: ||  :   | : :|  |    |   | | || |::|||||| |
Db   1081 NGQLRSILVKDTQAMKEMHFHPKALKDVKGQIGAPMPGKVIDIKVAAGDKVAKGQPLCVL 1140

Qy   1120 EAMKMEATITASVDGKIERVVVPAATKVEGGDLIVVV 1156
           ||||| | |||||| ||:  | : : ||| ||:| :
Db   1141 SAMKMETVVTSPMEGTIRKVHVTKDMTLEGDDLILEI 1177
```

```
RESULT 2
PYC_RAT
ID   PYC_RAT         STANDARD;       PRT;   1178 AA.
AC   P52873; Q64555;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic
DE   carboxylase) (PCB).
GN   PC.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=96096548; PubMed=8522203;
RA   Lehn D.A., Moran S.M., Macdonald M.J.;
RT   "The sequence of the rat pyruvate carboxylase-encoding cDNA.";
RL   Gene 165:331-332(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Liver;
RX   MEDLINE=96257760; PubMed=8687410;
RA   Jitrapakdee S., Booker G.W., Cassady A.I., Wallace J.C.;
RT   "Cloning, sequencing and expression of rat liver pyruvate
RT   carboxylase.";
RL   Biochem. J. 316:631-637(1996).
CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
CC       CARBOXYL GROUP TO PYRUVATE IN THE SECOND. CATALYZES IN A TISSUE
CC       SPECIFIC MANNER, THE INITIAL REACTIONS OF GLUCOSE (LIVER, KIDNEY)
CC       AND LIPID (ADIPOSE TISSUE, LIVER, BRAIN) SYNTHESIS FROM PYRUVATE.
CC   -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
```

```
CC        oxaloacetate
CC   -!- COFACTOR: BIOTIN AND MANGANESE (BY SIMILARITY).
CC   -!- PATHWAY: GLUCONEOGENESIS AND LIPOGENESIS.
CC   -!- SUBUNIT: Homotetramer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC       AND CARBAMYL PHOSPHATE SYNTHETASES.
CC   --------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------
DR   EMBL; U32314; AAA96256.1; -.
DR   EMBL; U36585; AAC52666.1; -.
DR   PIR; S68252; JC4391.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
KW   Ligase; Multifunctional enzyme; Biotin; Manganese; Gluconeogenesis;
KW   ATP-binding; Mitochondrion; Lipid synthesis; Transit peptide.
FT   TRANSIT       1     20       MITOCHONDRION (POTENTIAL).
FT   CHAIN        21   1178       PYRUVATE CARBOXYLASE.
FT   DOMAIN       21    549       BIOTIN CARBOXYLASE (BY SIMILARITY).
FT   DOMAIN      550   1000       CARBOXYLTRANSFERASE (BY SIMILARITY).
FT   DOMAIN     1096   1178       BIOTIN CARBOXYL CARRIER PROTEIN
FT                                (BY SIMILARITY).
FT   NP_BIND     198    203       ATP (BY SIMILARITY).
FT   ACT_SITE    328    328       BY SIMILARITY.
FT   BINDING    1144   1144       BIOTIN (BY SIMILARITY).
FT   CONFLICT    222    222       P -> S (IN REF. 2).
FT   CONFLICT    222    222       D -> I (IN REF. 2).
FT   CONFLICT    866    866       G -> R (IN REF. 2).
FT   CONFLICT    977    977
SQ   SEQUENCE   1178 AA;  129689 MW;  8E5FA19BC132A8DD CRC64;

Query Match             43.2%;   Score 2534.5;  DB 1;  Length 1178;
Best Local Similarity   46.2%;   Pred. No. 6.3e-126;
Matches  544;  Conservative 188;  Mismatches 409;  Indels 37;  Gaps 14;

Qy     7 GGLLLKGIITLVSTHTSSTLPA----------FKKILVANRGEIAVRAFRAALETGAA    54
Db     9 GGLRLLGV------RRSSTAFVASPNVRRLEYKPIKKVMVANRGEIAIRVFRACTELGIR   62

Qy    55 TVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLS  114
Db    63 TVAVYSEQDTGQMHRQKADEAYLIGRGLAPVQAYLHIPDIIKVAKENGVDAVHPGYGFLS  122

Qy   115 ENAQLARECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVL-AESTPSKNIDDIVK  173
Db   123 ERADFAQACQDAGVRFIGPSPEVVRKMGDKVEARAIAIAAGVPVVPGTNSPINSLHEAHE  182

Qy   174 SAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKLATEASREAEAAFGDGSVYVERAVINP  233
Db   183 FSNTYGFPIIFKAAYGGGGRGMRVVHSYEELEENYTRAYPEALAAFGNGALFVEKFIEKP  242

Qy   234 QHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSI  293
Db   243 RHIEVQILGDQYGNILHLYERDCSIQRRHQKVVEIAPATHLDPQLRSRLTSDSVKLAKQV  302

Qy   294 GYQGAGTVEFLVDEKGNHVFIEMNPIQVEHTVTEEVTEVDLVKAQMRLAAGATKELGL    353
Db   303 GYENAGTVEFLVKHGKHYFIEVNSRLQVEHTVTEEITDVDLVHAQIHVSEGRSLPDLGL   362

Qy   354 TQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQL-GGEITAHF  412
Db   363 RQENIRINGCAIQCRVTTEDPARSFQPDTGRIEVFRSGEGMGIRLDNASAFQGAVISPHY  422

Qy   413 DSMLVKMTCRGSDFETAVARAQRAIAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIG  472
Db   423 DSLLVKVIAHGKDHPTAATKMSRALAEFRVRGVKTNIPFLQNVLNNQQFLAGIVDTQFID  482

Qy   473 DHPHLLQAPPADDEQGRILDYLADVTVNKPHGVRPKDVA-APIDKL-PNIKDLPLPRGSR  530
Db   483 ENPELFQLRPAQNRAQKLLHYLHYLGHVMVNGPTTPIPVKVSPSPVDFIVPVVPIGPPAGFR  542

Qy   531 DRLKQLGPAAFARDLREQDALAVTDTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELL   590
Db   543 DILLREGPEGFARAVRNHQGLLLMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFNNLF  602

Qy   591 SVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAF  650
Db   603 SIENWGGATFDVAMRFLYECPWRRLQELRELIPNIPFQMLLRGANAVGYTNYPDNVVEKF  662

Qy   651 VKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETNTAVEVAMAYSGDLSDPNEKLYTLD   710
Db   663 CEVAKENGMDVFRIFDSLNYLPNMLLGMEAAGSAG-GVVEAAISYTGDVVADPSRTKYSLE  721

Qy   711 YYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVHTHDTAGGQLA  769
Db   722 YYMGLAEELVRAGTHILCIKDMAGLLKPAACTMLVSSLRDRFPDLPLHIHTHDTSGSGVA  781

Qy   770 TYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAV  829
Db   782 AMLACAQAGADVVDVAVDSMSGMTSQPSMGALVACTKGTPLDTEVPLERVFDYSEYWEGA  841

Qy   830 RGLLVLPFESGTPGPTGR--VYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEML  887
Db   842 RGLYAAFDCTATMKSGNSDVYENEDPGGQYTNLHFQAHSMGLGSKFKEVKKAYYEANQML  901

Qy   888 GRPTKVTPSSKVVGDLALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPE  947
Db   902 GDLIKVTPSSKIVGDLAQPFMVQNGLSRAEEAQAEELSFPRSVVEFLQGYIGIPHGGFPE  961

Qy   948 PLRTRALEG--RSEGKAPLTEVPEEQAHLDAD------DSKERRNSLNRLLFPKPTEEF   999
Db   962 PFRSKVLKDLPRIEG-GPGASLPPLNLKELEKDLIDRHGEEVTPEDVLSAAMPDVFAQF  1020

Qy  1000 LEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVAN 1059
Db  1021 KDFTATFGPLDSLNTRLFLQGPKIAEEFFEVELERGKT-LHIKALAVSDLNRAGGRQVFFE 1079

Qy  1060 VNGQIRPMVRDRSVESVTATAEKADSNKGHVAAPFAG-VVTVAEGDEVKAGDAVAI    1118
Db  1080 LNGQLRSIIVKDTQAMKEMHFHPKALKDVKGQIGAPMPGKVIDVKVAAGAKVVKGQPLCV 1139

Qy  1119 IEAMKMEATITASVDGKIERVVPAATKVEGGDLIVVV                        1156
Db  1140 LSAMKMETVVTSPMEGTIRKVHVTKDMTLEGDDLILEI                       1177

RESULT 3
PYC_HUMAN
ID   PYC_HUMAN      STANDARD;      PRT;   1178 AA.
AC   P11498; Q16705;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic
DE   carboxylase) (PCB).
```

```
GN   PC.
OS   Homo sapiens (Human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney, and Liver;
RX   MEDLINE=95002202; PubMed=7918683;
RA   Wexler I.D., Du Y., Lisgaris M.V., Mandal S.K., Freytag S.O.,
RA   Yang B.-S., Liu T.-C., Kwon M., Patel M.S., Kerr D.S.;
RT   "Primary amino acid sequence and structure of human pyruvate
RT   carboxylase.";
RL   Biochim. Biophys. Acta 1227:46-52(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=94324922; PubMed=8048912;
RA   MacKay N., Rigat B., Douglas C., Chen H.S., Robinson B.H.;
RT   "cDNA cloning of human kidney pyruvate carboxylase.";
RL   Biochem. Biophys. Res. Commun. 202:1009-1014(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney, and Liver;
RA   Walker M.E., Jitrapakdee S., Val D.L., Wallace J.C.;
RL   Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 1083-1178 FROM N.A.
RX   MEDLINE=87212051; PubMed=3555348;
RA   Lamhonwah A.-M., Quan F., Gravel R.A.;
RT   "Sequence homology around the biotin-binding site of human propionyl-
RT   CoA carboxylase and pyruvate carboxylase.";
RL   Arch. Biochem. Biophys. 254:631-636(1987).
RN   [6]
RP   SEQUENCE OF 1135-1178 FROM N.A.
RX   MEDLINE=85030380; PubMed=6548474;
RA   Freytag S.O. Collier K.J.;
RT   "Molecular cloning of a cDNA for human pyruvate carboxylase.
RT   Structural relationship to other biotin-containing carboxylases and
RT   regulation of mRNA content in differentiating preadipocytes.";
RL   J. Biol. Chem. 259:12831-12837(1984).
RN   [7]
RP   VARIANTS PC DEFICIENCY THR-610 AND ILE-743.
RX   MEDLINE=98254451; PubMed=9585612;
RA   Carbone M.A., MacKay N., Ling M., Cole D.E.C., Douglas C., Rigat B.,
RA   Feigenbaum A., Clarke J.T.R., Haworth J.C., Greenberg C.R.,
RA   Seargeant L., Robinson B.H.;
RT   "Amerindian pyruvate carboxylase deficiency is associated with two
RT   distinct missense mutations.";
```

```
RL   Am. J. Hum. Genet. 62:1312-1319(1998).
RN   [8]
RP   VARIANTS PC DEFICIENCY ALA-145 AND CYS-451.
RX   MEDLINE=98244401; PubMed=9585002;
RA   Wexler I.D., Kerr D.S., Du Y., Kaung M.M., Stephenson W., Lusk M.M.,
RA   Wappner R.S., Higgins J.J.;
RT   "Molecular characterization of pyruvate carboxylase deficiency in two
RT   consanguineous families.";
RL   Pediatr. Res. 43:579-584(1998).
CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
CC       CARBOXYL GROUP TO PYRUVATE IN THE SECOND. CATALYZES IN A TISSUE
CC       SPECIFIC MANNER, THE INITIAL REACTIONS OF GLUCOSE (LIVER, KIDNEY)
CC       AND LIPID (ADIPOSE TISSUE, LIVER, BRAIN) SYNTHESIS FROM PYRUVATE.
CC   -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
CC       oxaloacetate.
CC   -!- COFACTOR: BIOTIN AND MANGANESE.
CC   -!- PATHWAY: GLUCONEOGENESIS AND LIPOGENESIS.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC   -!- DISEASE: Defects in PC are the cause of pyruvate carboxylase
CC       deficiency (PC deficiency) [MIM:266150] leading to lactic
CC       acidosis, mental retardation and death. It occurs in three forms:
CC       mild or type A, severe neonatal or type B, and a very mild
CC       lacticacidemia.
CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC       AND CARBAMYL PHOSPHATE SYNTHETASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U04641; AAA99537.1; -.
DR   EMBL; S72370; AAB31500.1; -.
DR   EMBL; U30891; AAA82937.1; -.
DR   EMBL; BC011617; AAH11617.1; -.
DR   EMBL; M26122; AAA36423.1; -.
DR   EMBL; K02282; AAA60033.1; -.
DR   PIR; G01933; JC2460.
DR   HSSP; P24182; 1BNC.
DR   Genew; HGNC:8636; PC.
DR   GK; P11498; -.
DR   MIM; 266150; -.
DR   GO; GO:0005524; F:ATP binding activity; TAS.
DR   GO; GO:0009374; F:biotin binding activity; TAS.
DR   GO; GO:0004736; F:pyruvate carboxylase activity; TAS.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; Pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
KW   Ligase; Multifunctional enzyme; Biotin; Manganese; Gluconeogenesis;
KW   ATP-binding; Mitochondrion; Lipid synthesis; Transit peptide;
KW   Disease mutation.
FT   TRANSIT       1     20       MITOCHONDRION (POTENTIAL).
FT   CHAIN        21   1178       PYRUVATE CARBOXYLASE.
FT   DOMAIN       21    549       BIOTIN CARBOXYLASE (BY SIMILARITY).
```

```
FT   DOMAIN        550    1000       CARBOXYLTRANSFERASE (BY SIMILARITY).
FT   DOMAIN       1096    1178       BIOTIN CARBOXYL CARRIER PROTEIN
FT                                   (BY SIMILARITY).
FT   NP_BIND       198     203       ATP (BY SIMILARITY).
FT   ACT_SITE      328     328       BY SIMILARITY.
FT   BINDING      1144    1144       BIOTIN (BY SIMILARITY).
FT   VARIANT       145     145       V -> A (in PC deficiency; mild).
FT                                   /FTId=VAR_015199.
FT   VARIANT       451     451       R -> C (in PC deficiency; mild).
FT                                   /FTId=VAR_015200.
FT   VARIANT       610     610       A -> T (in PC deficiency; mild).
FT                                   /FTId=VAR_008095.
FT   VARIANT       743     743       M -> I (in PC deficiency; mild).
FT                                   /FTId=VAR_008096.
FT   CONFLICT      225     226       LA -> WP (IN REF. 2).
FT   CONFLICT      352     352       A -> S (IN REF. 3).
FT   CONFLICT      385     386       RS -> PT (IN REF. 2).
FT   CONFLICT      486     487       EL -> DV (IN REF. 2).
FT   CONFLICT      638     638       P -> R (IN REF. 2).
FT   CONFLICT      729     729       E -> A (IN REF. 2).
FT   CONFLICT      774     775       DT -> AP (IN REF. 2).
SQ   SEQUENCE     1178 AA;   129633 MW;  381F527553A20095 CRC64;

Query Match             43.2%;    Score 2531.5;  DB 1;  Length 1178;
Best Local Similarity   46.0%;    Pred. No. 9e-126;
Matches   541;   Conservative 186;   Mismatches 415;   Indels 35;   Gaps 13;

Qy      7  GGLLLKGIITILVSTHTSSTLPA--------------FKKILVANRGEIAVRAFRAALETGAA  54
                                         :|:||:||.||||||:||||:||||||||
Db      9  GGLRLLGI------RRTSTAPAASPNVRRLEYKPIKKVMVANRGEIAIRVFRACTELGIR    62

Qy     55  TVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPYGFLS   114
           ||||      :||| |||  :||:||||||||:|||| |||||| |||||||||||||
Db     63  TVAIYSEQDTGQMHRQKADEAYLIGRGLAPVQAYLHIPDIIKVAKENNVDAVHPGYGFLS  122

Qy    115  ENAQLARECAENGITFIGPTPEVLDITGDKSRAVTAAKKAGLPVL-AESTPSKNIDDIVK  173
           |||| ||| |||  |:||| |||||:||||:|||:|||||| ||: :|||:|||::|||
Db    123  ERADFAQACQDAGVRFIGPSPEVVRKMGDKVEARAIAIAAGVPVVPGTDAPITSLHEAHE  182

Qy    174  SAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKLATEASREAEAAFGDGSVYVERAVINP  233
           ||| |:| ||||||:|||||||||||| |||:|||:||:|:|| ||||:||||||||||
Db    183  FSNTYGFPIIFKAAVGGGGRGMRVVHSYEELEENYTRAYSEALAAFGNGALFVEKFIEKP  242

Qy    234  QHIEVQILGDRTGEVVHLYERDCSLQRRHQKVEIAPAQHLDPELRDRICADAVKFCRSI  293
           :||||||||||: ||||||||||||||||||||:|||||||||||||:|||||||:||
Db    243  RHIEVQILGDQYGNILHLYERDCSIQRRHQKVEIAPAAHLDPQLRTRLTSDSVKLAKQV  302

Qy    294  GYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGL  353
           ||| ||||||:||: || :||||||||||||:|:|||:||||||||:||||| ||:| |
Db    303  GYENAGTVEFLVDRHGKHYFIEVNSRLQVEHTVTEITDYDLVHAQIHVAEGRSLPDLGL  362

Qy    354  TQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQL-GGEITAHF  412
           :: | |||| |::|:||:|||:::|:|  |||:|||: :|| :|| | |: ||::||||
Db    363  RQENIRINGCAIQCRVTTEDPARSFQPDTGRIEVFRSGEGMGIRLDNASAFQGAVISPHY  422

Qy    413  DSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIG  472
           |||:|:|:  | | | || |  :||||||||||||||||:||||||||||||||||||
Db    423  DSLLVKVIAHGKDHPTAATKMSRALAEFRVRGVKTNIAFLQNVLNNQQFLAGTVDTQFID  482

Qy    473  DHPHLLQAPPADDEQGRILDYLADVTVNKPHGVRP-KDVAAPIDK-LPNIKDLPLPRGSR  530
           :: |  |  || || ||:||||: ||:||||:|| ||| ||::| |::||||||:|||
Db    483  ENPELFQLRPAQNRAQKLLHYLGHVMVNGPTTPIPVKASPSPTDPVVPAVPIGPPPAGFR  542

Qy    531  DRLKQLGPAAFARDILREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELL  590
           || :|| | | |:  || | || |||||||||||:||:|||:||:||| ||| || ||
Db    543  DILLREGPEGFARAVRNHPGLLLMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFSKLF  602

Qy    591  SVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVIQMLLRGRNTVGTTPYPDSVCRAF  650
           |: || || ||||||||:|:| || :|:|||| ||:||:|||||| ||||:|||| |
Db    603  SMENWGGATFDVAMRFLYECPWRRLQELRELIPNIPFQMLLRGANAVGYTNYPDNVVFKF  662

Qy    651  VKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLD  710
           :  |  |||::|| :|: :: |         |   | | |       |: | :| ||
```

```
Db    663  CEVAKENGMDVFRVFDSLNVLPNMLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQ  721

Qy    711  YYLKMAEEIVKSGAHIIAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVTHDTAGGQLA  769
           ||: :|||:|| | :||:|||||||:|:|| || : |:|:||  ||| |:||:|:|| |
Db    722  YYMGLAEELVRAGTHILCIKDMAGLLKPTACTMLVSSLRDRFPDLPLHTHDTSGAGVA   781

Qy    770  TYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAV  829
           | ||||:||| ||| |:| ||| ||||:::||||:|:   |:|  : :|||||:|||:|
Db    782  AMLACAQAGADVVDVAADSMSGMTSQPSMGAIVACTRGTPLDTEVPMERVFDYSEYWEGA  841

Qy    830  RGLYLPFESGTPGPTGR--VYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEML  887
           |||| | :|: ||||   :| | | |||| ||  | | || |  ::| |::|  | ||:
Db    842  RGLIYAAFDCTATMKSGNSDVYENEIPGGQYTNLHFQAHSMGLGSKFKEVKKAYYEANQML  901

Qy    888  GRPTKVTPSSKVVGDLALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPE  947
           || |||||||:|:||| |:||||||  |:|   :|| |:||||  ||:|| |||:| |
Db    902  GDLIKVTPSSKIVGDLAQFMVQNGLSRAEAEAQAEELSFPRSVVEFLQGYIGVPHGGFPE  961

Qy    948  PLRTRAL------EGRSEGKAPLTEVPEEQAHLDA-DDSKERRNSLNRLFLFPKPTEEFL  1000
           ||:|:|        ||:|  | ||| : :|| | | ::|||::|::|:|:| |||:| |
Db    962  PFRSKVLKDLPRVEGRPGASLPPLDLQALEKELVDRHGEEVTPEDVLSAAMYPDVFAHFK  1021

Qy   1001  EHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANV  1060
           :|:: ||||:|  | :||:|::| ::|: ||: ||| |:|  | :||| |||||:: |
Db   1022  DFTATFGPLDSLNTRLFLQGPKIAEEFEVELERGKI-LHIKALAVSDLNRAGQRQVFFEL  1080

Qy   1061  NGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAII  1119
           |||:| |||  | |:| | | |||:  ||:||||| |  :|::: |||: |||| | ||
Db   1081  NGQIRSILVKDTQAMKEMHFHPKALKDYKGQIGAPMPGKVIDIKVVAGAKVAKGQPLCVL  1140

Qy   1120  EAMKMEATITASVDGKIERVVVPAATKVEGGDLIVVV  1156
           ||||||: |:|| ||| |||:|:|  |:|| | |:|
Db   1141  SAMKMETVTISPMEGTVRKVHVTKDMTLEGDDLILEI  1177
```

```
RESULT 4
PYC1_YEAST
ID   PYC1_YEAST     STANDARD;     PRT;   1178 AA.
AC   P11154;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyruvate carboxylase 1 (EC 6.4.1.1) (Pyruvic carboxylase 1) (PCB 1).
GN   PYC1 OR PYV OR YGL062W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=88298805; PubMed=3042770;
RA   Lim F., Morris C.P., Occhiodoro F., Wallace J.C.;
RT   "Sequence and domain structure of yeast pyruvate carboxylase.";
RL   J. Biol. Chem. 263:11493-11497(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=97377993; PubMed=9234674;
RA   Feuermann M., de Montigny J., Potier S., Souciet J.-L.;
RT   "The characterization of two new clusters of duplicated genes
RT   suggests a 'Lego' organization of the yeast Saccharomyces cerevisiae
RT   chromosomes.";
RL   Yeast 13:861-869(1997).
RN   [3]
RP   SEQUENCE OF 1003-1178 FROM N.A.
RX   MEDLINE=87241529; PubMed=3036126;
RA   Morris C.P., Lim F., Wallace J.C.;
RT   "Yeast pyruvate carboxylase: gene isolation.";
RL   Biochem. Biophys. Res. Commun. 145:390-396(1987).
CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
CC       CARBOXYL GROUP TO PYRUVATE IN THE SECOND.
```

```
CC   -!- CATALYTIC ACTIVITY: ATP + Pyruvate + HCO(3)(-) = ADP + phosphate +
CC       oxaloacetate.
CC   -!- COFACTOR: BIOTIN AND ZINC.
CC   -!- PATHWAY: Gluconeogenesis.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC       AND CARBAMYL PHOSPHATE SYNTHETASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J03889; AAA34843.1; -.
DR   EMBL; Z72584; CAA96765.1; -.
DR   PIR; S64066; QYBYP.
DR   HSSP; P24182; 1BNC.
DR   SGD; S0003030; PYC1.
DR   GO; GO:0005829; C:cytosol; IDA.
DR   GO; GO:0004736; F:pyruvate carboxylase activity; IDA.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; Biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Ligase; Multifunctional enzyme; Biotin; Gluconeogenesis; ATP-binding;
KW   Zinc; Multigene family.
FT   NP_BIND     182    187       ATP (POTENTIAL).
FT   ACT_SITE    312    312       BY SIMILARITY.
FT   BINDING    1135   1135       BIOTIN (BY SIMILARITY).
FT   CONFLICT    462    462       T -> G (IN REF. 1).
FT   CONFLICT    493    493       V -> D (IN REF. 1).
FT   CONFLICT    595    595       R -> A (IN REF. 1).
FT   CONFLICT    619    619       E -> Q (IN REF. 1).
FT   CONFLICT    664    664       G -> S (IN REF. 1).
FT   CONFLICT    772    772       A -> R (IN REF. 1).
FT   CONFLICT    879    879       E -> Q (IN REF. 1).
FT   CONFLICT    909    909       Q -> K (IN REF. 1).
SQ   SEQUENCE   1178 AA;  130099 MW;  BC7110A8AFB23E04 CRC64;

Query Match             41.9%;  Score 2457;  DB 1;  Length 1178;
Best Local Similarity   46.6%;  Pred. No. 7.5e-122;
Matches  540;  Conservative  175;  Mismatches  400;  Indels  44;  Gaps  16;

Qy    31 KILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEG--SPVKAY 88
          |||||||||| | | || |  |  |||||  |||  ::  | |   |  | |    |||| |
Db    21 KILVANRGEIPIRIFRTAHELSMQTVAIYSHEDRLSTHKQKADEAYVIGEVGVGYGQYTPVGAY 80

Qy    89 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 148
          | |||| :||:: | : | ||||||||| : | : || | ||| |||:| :||||:|| 
Db    81 LAIDEIISIAQKHQVDFIHPGYGFLSENSEFADKVVKAGITWIGPPAEVIDSVGDKVSAR 140

Qy   149 TAAKKAGLPVLAEST-PSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKL 207
          ::| :||:|.: | : : | :| | | ||| |||:| |||.:||||||||:|.  |  : :|
Db   141 NLAAAKANVPTVPGTPGPIETVEEAALDFVNEYGYPVIIKAAFGGGGRGMRVVREGDDVADA 200
```

```
Qy   208 ATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVE 267
          :|| || | ||| |:|.|||| : :| :|||||:| | .: |||::||||.||||||||| 
Db   201 FQRATSEARTAFGNGTCFVERFLDKPKHIEVQLLADNHGNVVHLFERDCSVQRRHQKVVE 260

Qy   268 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 327
          :||| .| | ::||| .||| | : ||: .|||.||||| : : : |||:|||||||||:| 
Db   261 VAPAKTLPREVRDAILTDAVKLAKECGYRNAGTAEFLVDNQNRHYFIEINPRIQVEHTIT 320

Qy   328 EEVTEVDLVKAQMRLAAGATLKELGLTGDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 387
          ||: : ||:||| : ||||:|: |||: |||  ||.|:||||||||| : | |||| |. 
Db   321 EEIIGIDIVAAQIQIAAGASLPQLGLFQDKITTRGFAIQCRITTEDPAKNFQPDTGRIEV 380

Qy   388 YRSPGGAGVRLDGA-AQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVA 446
          |||.||.|||||.: |: || |::.:||||||.:| | :   :. .:: | |:| : ||. 
Db   381 YRSAGGNGVRLDGGNAYAGTIISPHYDSMLVKCSCSGSTYEIVRKMIRALIEFRIRGVK 440

Qy   447 TNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVN---KP 502
          |||.|| :||.: ||:  | .| |: : | | : :|  : :: :| ||||| ::   | 
Db   441 TNIPFLLTLLTNPVFIEGTYWTTFIDDTPQLFQMVSSQNRAQKLLHYLADVAVNGSSIKG 500

Qy   503 HGVRPKDVAAPIDKLPNIKDL--------PLPRGSRDRLKQLGPAAFARDLREQDALA 552
           :| : ::.:  : :: |  |         |: | |: : : | || || ||  | .:|
Db   501 QIGLPKLKSNP--SVPHLHDAQGNVINVTKSAAPPSGWRQVLLEKGPAEFARQVRQFNGTL 558

Qy   553 VTDTTFRDAHQSLLATRVRSFALKPAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPW 612
          : |||:|||||||||||||: : |.|  :| .||:|| | || | |:||||||| ||||
Db   559 LMDTTWRDAHQSLLATRVRTHDLATIAPTTAHALAGRFALECWGGATFDVAMRFLHEDPW 618

Qy   613 DRLDELRERAMPNVNIQMLLRGRNTVGYTPYPPDSVCRAPVKEAASSGVDIFRIFDALNDVS 672
          :||: || ..|||.:.|||||| |  | : :|:|:   : || :: || |||||:||||:.
Db   619 ERLRKLRSLVPNIPFQMLLRGANGVAYSSLPDNAIDHFVKQAKDNGVDIFRVFDALNDLE 678

Qy   673 QMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLKMAEEIVKSGAHILAIKDM 732
          |:: :.||| :.  ||| |..:|||: : ..:: | |||:::: ::||: |::|.|||||
Db   679 QLKVGVDAVKKAG-GVVEATVCFSGDMLQPGKK-YNLDYYLEIIAEKIVQMGTHILGIKDM 736

Qy   733 AGLLRPAAVTKLVTALRREF-DLPVHVHTHDTRAGGQLATYFAAAQAGADAVDGASAPLSG 791
          || ::|||..:|: .|| :: ||| ||| ||:||:.:::.| ::|.|||| ||.::. ||
Db   737 AGTMKPAAAKLLIGSLRAKYPDLPIHVHHDSAGTAVASMTACALAGADVVDVAINSMSG 796

Qy   792 TTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHE 851
          :||||| :|::.:: :  : | :  :||.::: ||:|:| || .||. :  |  | |:|| 
Db   797 LTSQPSINALLASL-EGNIDTGINVEHVRELDAYWAEMRLLYSCFEADLKGPDPEVYQHE 855

Qy   852 IPGGQLSNLRAQATALGLADRFELIEDNVAAVNEMLGRPTKVTPSSKVVGDLALHLVGAG 911
          ||||||:||  ||  ||| : :.  ..::.|.:: ||.: ||| :||||||| : :|.. 
Db   856 IPGGQLTNLLFQAQQLGLGEQWAETKRAYREANYLLGDIVKVTPTSKVVGDLAQFMVSNK 915

Qy   912 VDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEE 971
          : ::|:| |  : ::|||:: ::.| :| .| |::||:: : : . :   ::|| .| ||
Db   916 LTSDDVRRLANSLDFPDSVMDFFEGLIGQPYGYGGFPEPFRSDVLRNK---RRKLTCRPGLE 972

Qy   972 QAHLDADDSKERRNSLNRL---------LFPKPTEEFLEHRRRFGNTSALDDREFFY 1019
          :.  ||||::: :|||:::         |:|| ||||:||||| |||||||||||||
Db   973 LEPFDLE--KIREDLQNRFGDVDECDVASYNMYPRVYEDFQKMRETYGDLSVLPTRSFLS 1030

Qy  1020 GLVEGRETLIRLPDVRTPLLVRLDAISEPDDK-GMRNVVANVNGQIRPMRVRDRSVESVT 1078
           | :  |: :.:: || ||::: || : | :: |:: |:: :||::|::||.|||.: ::
Db  1031 PLETDEEIEVVIEQGKT-LIIKLQAVGDLNKKTGEREVYFDLNGEMRKIRVADRSQKVET 1089

Qy  1079 AVAEKADSSNKGHVAAPFAGV-VTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIE 1137
          .|:.| ||.::| :.|||:||: : | |: | :::|| :|:: :||||| |::|.||::
Db  1090 VTKSKADMHDPLHIGAPMAGVIVEVKVHKGSLIKKGQPVAVLSAMKMEMIISSPSDGQVK 1149

Qy  1138 RVVPAATKVEGGDLIVVV 1156
          .||.:.. |:|.||:|:|
Db  1150 EVFVSDGENVDSDLLVLL 1168

RESULT 5
PYC2_YEAST
ID   PYC2_YEAST     STANDARD;      PRT;   1180 AA.
AC   P32327;
```

```
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyruvate carboxylase 2 (EC 6.4.1.1) (Pyruvic carboxylase 2) (PCB 2).
GN   PYC2 OR YBR218C OR YBR1507.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX   MEDLINE=92017667; PubMed=1921979;
RA   Stucka R., Dequin S., Salmon J.-M., Gancedo C.;
RT   "DNA sequences in chromosomes II and VII code for pyruvate
RT   carboxylase isoenzymes in Saccharomyces cerevisiae: analysis of
RT   pyruvate carboxylase-deficient strains.";
RL   Mol. Gen. Genet. 229:307-315(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RA   Dubois E., el Bakkoury M., Glansdorff N., Messenguy F., Pierard A.,
RA   Scherens B., Vierendeels F.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96128067; PubMed=8554526;
RA   Val D.L., Chapman-Smith A., Walker M.E., Cronan J.E. Jr.,
RA   Wallace J.C.;
RT   "Polymorphism of the yeast pyruvate carboxylase 2 gene and protein:
RT   effects on protein biotinylation.";
RL   Biochem. J. 312:817-825(1995).
CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
CC       CARBOXYL GROUP TO PYRUVATE IN THE SECOND.
CC   -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
CC       oxaloacetate.
CC   -!- COFACTOR: BIOTIN AND ZINC.
CC   -!- PATHWAY: Gluconeogenesis.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- INDUCTION: By glucose.
CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC       AND CARBAMYL PHOSPHATE SYNTHETASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X59890; CAA42544.1; -.
DR   EMBL; Z36067; CAA85182.1; -.
DR   EMBL; U35647; AAC49147.1; -.
DR   PIR; S46094; S46094.
DR   HSSP; P24182; 1BNC.
DR   SGD; S000422; PYC2.
DR   GO; GO:0005829; C:cytosol; IDA.
DR   GO; GO:0004736; F:pyruvate carboxylase activity; IDA.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; Biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
```

```
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Ligase; Multifunctional enzyme; Biotin; Gluconeogenesis; ATP-binding;
KW   Zinc; Multigene family.
FT   NP_BIND     183    188       ATP (POTENTIAL).
FT   ACT_SITE    313    313       BY SIMILARITY.
FT   BINDING    1136   1136       BIOTIN (BY SIMILARITY).
FT   CONFLICT     15     15       S -> C (IN REF. 1).
FT   CONFLICT    132    132       D -> E (IN REF. 1).
FT   CONFLICT    238    238       N -> K (IN REF. 1).
FT   CONFLICT    268    268       L -> Y (IN REF. 1).
FT   CONFLICT    546    546       S -> C (IN REF. 1).
FT   CONFLICT    642    642       N -> T (IN REF. 1).
FT   CONFLICT    771    773       GTA -> STR (IN REF. 1).
FT   CONFLICT    831    831       W -> R (IN REF. 1).
FT   CONFLICT    839    839       S -> P (IN REF. 1).
FT   CONFLICT   1001   1001       Y -> N (IN REF. 1).
FT   CONFLICT   1155   1155       K -> R (IN REF. 1).
FT   CONFLICT   1178   1178       Q -> P (IN REF. 1).
FT   CONFLICT   1180   1180       K -> KVIFTR (IN REF. 1).
SQ   SEQUENCE  1180 AA;  130166 MW;  AD60DA3A60F5E001 CRC64;
```

```
Query Match             41.5%;  Score 2436;  DB 1;  Length 1180;
Best Local Similarity   45.5%;  Pred. No. 9.6e-121;
Matches  530;  Conservative  185;  Mismatches  407;  Indels  44;  Gaps  16;

Qy    24  STLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEG-  82
              |||||||||||||: | ::|:||:|  ||:||| ||| ||| | ||
Db    15  SLLGEKNKILVANRGEIPIRIFRSAHELSMRTIAIYSHEDRLSMHRLKADEAYVIGEGQ  74

Qy    83  -SPVKAYLDIDEIIGAAKKVKADAITPGYGFLSENAQLARECAENGITFIGPTPEVLDLT  141
           :|| | |||:|||: ||:|  ||||||||||:||:|:|||:||:| :|| |||
Db    75  YTPVGAYLAMDEIIEIAKKHKVDFIHPGYGFLSENSEFADKVVKAGITWIGPPAEVIDSV  134

Qy   142  GDKSRAVTAAKKAGLPVLAEST-PSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSS  200
           ||| || | |||||||| :||  | :| ||||| : ||  |||||||  |||||||||  |
Db   135  GDKVSARHLAARANVPTVPGTPGPIETVQEALDFVNEYGYPVIIKAAFGGGGRGMRVVRE  194

Qy   201  PDELRKLATEASREAEAAFGDGSVYERAVINPQHIEVQILGDRTGEVVHLYERDCSLQR  260
            :: || || |||:||||:||:||  || :|||:| |||:|  | ||| :|||| || |
Db   195  GDDVADAFQRATSEARTAFGNGTCFVERFLDKPKHIEVQLLADNHGNVVHLFERDCSVQR  254

Qy   261  RHQKVVEIAPAQHLDPELRDRICADAVKFPCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRI  320
           |||||||||| || ||||||| |:||||||||| :|| |||:||| || |||||||||||
Db   255  RHQKVVEVAPAKTLPREVRDAILTDAVKLAKVCGYRNAGTAEFLVDNQNRHVFIEINPRI  314

Qy   321  QVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRP  380
           |||||||||||||||:| |:| |||||  |||||||||||||||| ||||||||| | ||
Db   315  QVEHTITEEITGIDIVSAQIQIAAGATLTQLGLLQDKITTRGFSIQCRITTEDPSKNFQP  374

Qy   381  DTGTTTAYRSPGGAGVRLDGA-AQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAE  439
           ||| | || ||| || |||||  || ||||:||||:||||||||| |||| || | |
Db   375  DTGRLEVYRSAGGNGVRLDGGNAYAGATISPHYDSMLVKCSCSGSTYEIVRRKMIRALIE  434

Qy   440  FTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTV  499
           | | |  | ||||| |:: |:|  ||      |||  ||:|  ||: | | |: | | |
Db   435  FRIRGVKTNIPFLLTLTNPVFIEGTYWTTFIDDTPQLFQMVSSQNRAQKLLHYLADLAV  494

Qy   500  N----KPHGVRPKDVAAPIDKLPNIKDL--------PLPRGSRDRLKQLGPAAPARDL  545
           |      | :  | || |    : |: |           ||:  | : || | |:|
Db   495  NGSSIKGQIGLPKLKSNP--SVPHLHDAQGNVINVTKSAPPSGWRQVLLEKGPSEFAKQV  552

Qy   546  REQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLIPELLSVEAWGGATYDVAMR  605
           ||:||| || | |||||:|||||||:|||  |||   |||  |  ||| ||||||| |||
Db   553  RQFNGTLLMDTTWRDAHQSLLATRVRTHDLATIAPTTAHALAGAFALECWGGATFDVAMR  612

Qy   606  FLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIF  665
           |||||||:||: | || :||:|| ||||||:|| |||||| |||| |||:|:|||||||
Db   613  FLHEDPWERLRKLRSLVPNIPFQMLLRGANGVAYSSLPDNAIDHFVKQAKDNGVDIFRVF  672
```

```
RESULT 6
PYC_PICPA
ID   PYC_PICPA      STANDARD;      PRT;   1189 AA.
AC   P78992;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Pyruvate carboxylase (EC 6.4.1.1) (Pyruvic carboxylase) (PCB).
GN   PYC1.
OS   Pichia pastoris (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Pichia.
OX   NCBI_TaxID=4922;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98301182; PubMed=9639311;
RA   Menendez J., Delgado J., Gancedo C.;
RT   "Isolation of the Pichia pastoris PYC1 gene encoding pyruvate
RT   carboxylase and identification of a suppressor of the Pyc
RT   phenotype.";
RL   Yeast 14:647-654(1998).
CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
CC       CARBOXYL GROUP TO PYRUVATE IN THE SECOND (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP + Pyruvate + HCO(3)(-) = ADP + phosphate +
CC       oxaloacetate.
CC   -!- COFACTOR: BIOTIN AND ZINC.
CC   -!- PATHWAY: Gluconeogenesis.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC       AND CARBAMYL PHOSPHATE SYNTHETASES.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
DR   EMBL; Y11106; CAA71993.1; -.
DR   HSSP; P24182; 1DV1.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L.
DR   InterPro; IPR005483; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   PRINTS; PR00098; CPSASE.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Ligase; Multifunctional enzyme; Biotin; Gluconeogenesis; ATP-binding;
KW   zinc.
FT   NP_BIND     185    190       ATP (POTENTIAL).
FT   ACT_SITE    315    315       BY SIMILARITY.
FT   BINDING    1140   1140       BIOTIN (BY SIMILARITY).
SQ   SEQUENCE   1189 AA;  131400 MW;  8B6E858079657914 CRC64;

Query Match             40.8%;  Score 2393;  DB 1;  Length 1189;
Best Local Similarity   45.5%;  Pred. No. 1.8e-118;
Matches   529;  Conservative 176;  Mismatches 419;  Indels 38;  Gaps 18;

Qy    23 SSTLPAFKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEG 82
          || |  ||||||||||:|||||| :||| ||  :|| ||| | |||  || : | |||
Db    16 SSLLGTMNKILVANRGEIPIRIFRTAHELSMNTVAIYSHEDRLSMHRLKADEAYVIGERG 75

Qy    83 --SPVKAYLDIDEIIGAAKKVKADAIYPGVGFLSENAQLARECAENGITFIGPTPEVLDL 140
           | | :||| ||| || |: | :|| :|||| ||| | | |  |||| :||| ||||
Db    76 QYSFVQAYLAIDEIIKIAVKHNVMIHPGYGFCSENSEFARKVEENGILWVGPSDTVIDA 135

Qy   141 TGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKS---AEGQTYPIFVKAVAGGGRGMRF 197
          :||  || |  || ||||  :|   :|| |::|      |:|  | ||| ::||||||
Db   136 VGDKVSARNLAYAANVPTV-PGTPGP-IEDVAQATAFVEEYGYPVIKAAFGGGGRGMRV 193

Qy   198 VSSPDELRLKATEASREARAAFGDGSVVYERAVINPQHIEVQILGDRTGEVVHLYERDCS 257
          ::|    |||  |::|||| |||  ||  ||| | |||  | |:  |||| ||| ||||:
Db   194 VREGDDIEDAFLRASSEAKTAFGNGTVFIRFLDKPKHIEVQLLADNYGNVIHLFERDCS 253

Qy   258 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQQAGTVEFLVDEKGNHVFIEMN 317
          |||||||| :|:|| ||||  ||:|:||||:||||||| |||||| ||  ||||||||:|
Db   254 VQRRHQKVARNCSAKTLPEVRNAILNDAVKLAKTANYRNAGTAEFLVDSQNRHYFIEIN 313

Qy   318 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATIKELGLTQDKIKTHGAALQCRITTEDPNNG 377
          |||||||| ||||:|:|||||| |||||| | :||| ||:|  || :| |||  || :|
Db   314 PRIQVEHTITEEITGVDIVAAQIQIAAGASLEQLGLLQEKITTRGFAIQCRITTEDPTKN 373

Qy   378 FRPDTGTITAYRSPGGAGVRLDGAAQLGDE-ITAHFDSMLVKMTCRGSDFETAVARAQRA 436
          |:|||| :|:|:|| |||:| |||  |||| :|:|||||:|:|| |  |  | || | |
Db   374 FQPDTGKIEVYRSSGNGVRLDGGNGFAGAVISPHYDSMLVKCSTSGSNYEIRRKMIRA 433

Qy   437 LAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRIILDYLAD 496
          |:|| :||| |||  |||: |||| |  | ||| |||||||||||  | :||| |||||
Db   434 LVEFRIRGVKTNIPFLLALLTHPVFMTSECWTTFIDDTPELFKILTSQNRAQKLLAYLGD 493

Qy   497 VTVN-------KPHGVRPKDVAAPIDKLPNIKD--LPLPRGSRDLRLKQLGPAAFAR 543
          ::|                : :|| || |:||   || |||||| || ||| | |
```

```
Qy   666 DALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAH 725
          ||||| || :|| | ||  |||| || | |||:|||||| |||:|| |||||::|| |
Db   673 DALNDLEQLKVGVNAVKKAG-GVVEATVCYSGDMLQPGKK-YNLDYYLEVVEKIVQMGTH 730

Qy   726 ILAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVTHDTAGGOLATYFAAAQAGADAVDG 784
          || |||||  | :|||| : |  || : :|||:|:|||:| |||| ||:||:| |||
Db   731 ILGIKDMAGTMKPAAAKLLIGSLRTKYPDLPIHVHSHDSAGTAVASMTACALAGADVVDV 790

Qy   785 ASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPT 844
          || | || :||:| | ::: |  |   ||| :| | |   | ||| || || | |
Db   791 AINSMSGLTSQPSINALLASL-EGNIDTGINVEHVRELDAYWAEMRLLYSCFEADLKGPD 849

Qy   845 GRVYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLA 904
          || ||||||| :| ||  | :|  ||  :|:   |   |  || ||| |:|||||||:|
Db   850 PEVYQHEIPGGQTNLLFQAQQLGLGEQWAETKRAYREANYLLGDIVKVTPTSKVVGDLA 909

Qy   905 LHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPL 964
          | | | | |    | || | |  |:  || |  |||  |||:   | |   |  |
Db   910 QFMVSNKLTSDDIRRLANSLDFDSVMDFFEGLIGQPYGGFPPELRSDVLRNK--RRKL 966

Qy   965 TEVPEEEQAHLDADDSKEPRRNSLNRL----------LFPKPTEEFLEHRRRFGNTSAL 1012
             : |  |  |  || |  |||   |          |||| ||| |||||:||| :|
Db   967 TCRGLELEPFDLE--KIREDLQNRFGDIDECDVASYNMYPRVYEDFQKIRETYGDLSVL 1024

Qy  1013 DDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDK-GMRNVANVNGQIRPMRYRD 1071
          ||  |||  | |  | :|  ||  | ::   |:::|   |||  | |:|  ||| |
Db  1025 PTKNFLAPAEPDEIEVTIEQKT-LIIKLQAVGDLNKKTGQREVYFELNGELRKIRVAD 1083

Qy  1072 RSVESVTATAEKADSSNKGHVAAPFAGV-VTVTVAEGDEVKAGDAVAIIEAMKMEATITA 1130
          || :|   |  | :   |  |  || |  :| |:| ||:|| | | | :|:||||  | :
Db  1084 KSQNIQSVAKPKADVHDTHQIGAPMAGVIIEVKVHKGSIVKKGESIAVLSAMKMEMVSS 1143

Qy  1131 SVDGKIERVVPAATKVEGGDLIVVV 1156
          ||||| | |  | | | | |:|:||
Db  1144 PADGQVKDVFIKDGESVDASDLLVL 1169
```

```
RT   jannaschii.";
RL   Science 273:1058-1073(1996).
RN   [2]
RP   SEQUENCE OF 1-12, AND FUNCTION.
RX   MEDLINE=21034791; PubMed=11195096;
RA   Mukhopadhyay B., Patel V.J., Wolfe R.S.;
RT   "A stable archaeal pyruvate carboxylase from the hyperthermophile
RT   Methanococcus jannaschii.";
RL   Arch. Microbiol. 174:406-414(2000).
CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE CARBOXYL
CC       GROUP TO PYRUVATE IN THE SECOND.
CC   -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
CC       oxaloacetate.
CC   -!- COFACTOR: ATP, MAGNESIUM (OR MANGANESE OR COBALT), PYRUVATE AND
CC       BICARBONATE.
CC   -!- ENZYME REGULATION: INHIBITED BY MAGNESIUM, WHEN ITS CONCENTRATION
CC       EXCEEDED THE ATP ONE, AND BY HIGH CONCENTRATION OF ATP AND ALPHA-
CC       KETOGLUTARATE.
CC   -!- PATHWAY: Gluconeogenesis.
CC   -!- SUBUNIT: HETEROOCTAMER OF FOUR A AND FOUR B SUBUNITS.
CC   -!- MASS SPECTROMETRY: MW=55500; METHOD=MALDI.
CC   -!- MISCELLANEOUS: ITS OPTIMUM PH IS 8.5 AND THE OPTIMUM TEMPERATURE
CC       IS 80-90 DEGREES CELSIUS.
CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC       AND CARBAMYL PHOSPHATE SYNTHETASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U67563; AAB99232.1; -.
DR   PIR; D64453; D64453.
DR   HSSP; P24182; 1BNC.
DR   TIGR; MJ1229; -.
DR   InterPro; IPR004549; AccC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR005483; CPase_L.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   PRINTS; PR00098; CPSASE.
DR   TIGRFAMs; TIGR00514; accC; 1.
DR   PROSITE; PS00866; CPSASE; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Ligase; Multifunctional enzyme; Gluconeogenesis; Magnesium; Pyruvate;
KW   ATP-binding; Complete proteome.
FT   NP_BIND       162    167       ATP (POTENTIAL).
FT   ACT_SITE      291    291       POTENTIAL.
SQ   SEQUENCE   501 AA;  55402 MW;  04D2E401892F872F CRC64;
```

```
Query Match             18.4%;  Score 1079;  DB 1;  Length 501;
Best Local Similarity   48.0%;  Pred. No. 8.5e-50;
Matches  217;  Conservative  85;  Mismatches  144;  Indels  6;  Gaps  4;

Qy     29 FKKIIVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 88
           |  |:||||||||  | :||||||||| |:|  |||| ||| || :
Db      2 FNKVLIANRGEIAIRIIRACWELGIKTVAVYSEADKRSLHATLADEAYCIG-PAPAAKSY 60
Qy     89 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTEVLLDTGDKSRAV 148
           | | |:|||||||||:||||||||:||||||| | | ||:||| :|
Db     61 LNIDAILNVAEKAKVDAIHPGYGFLAENAEFARAVKKAGPEFIGPNDAEAMGSKINAK 120
Qy    149 TAAKKAGLPVLAESTPS-KNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKL 207
           | :|||||| :|:     :|:|| ||||||| |:|||| |||||||||| :| ||||||
Db    121 KIMKKAGVPLIPGSEGAIEDIDEAIEIAEAIGFPVVVKASAGGGGMGMSVAYSKEELKEV 180
```

```
Db    494 LAVNGSSIKGQIGLPKLH--KEADIPSITDINGDVIDVSIPPPDGWRQFLLEKGPEQFAQ 551
Qy    544 DLREQDALAVTDTTFERDAHQSLLATRVRSFALKPAABAVAKLITPELLSVEAWGGATYDVA 603
           :|  :|| :| | ||||||||||||  | |  | : :|  | ||||||||||||:|
Db    552 QVRAFPGLMLMDTTWRDAHQSLLATRVRTHDLLNIAPATSYALHHAFALECWGGATFDVS 611
Qy    604 MRFLFEDPWDRIDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFR 663
           ||| ||| ||| :|: :|:: ||||| |: | ||| |||:|| ::|||:|
Db    612 MRFLHEDPWQRLRLRKAVPNIPFSMLLRGGNGVAYYSLPDNAIDHFLKQAKDTGVDVFR 671
Qy    664 IFDALNDVSQMRPAIDAVLETNIAVAEVAMAYSGDLSDPNEKLYTLDYIKMAEEIVKSG 723
           :|||: |: :|     : | ::  :  ||||:    :| | || : | :|||| :|
Db    672 VFDALNDIEQLKVGVDAVKKAG-GVVEATMCYSGDMLKPGKK-YNLEYYINLATEIVEMG 729
Qy    724 AHILAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVHTHDTAGGQLATYFAAAQAGADAV 782
           ||| :||||| ||| :  ||:|| |||:  :||:|||||| :| |  :| :| |||||
Db    730 THILAVKDMAGTLKPTAAKQLISALRRKFPSLPIHVHTHDSAGTGVASMVACARAGADVV 789
Qy    783 DGASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPPESGTPG 842
           | :| || |::|  |:|| :  | :    :   :||  : || : |
Db    790 TVRVNSMSCMTSQPSMSAFIASL-DGEIETGIPEANAREIDAYWAEMRLLYSCFEADLKG 848
Qy    843 PTGRVYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGD 902
           ||| :|  ||||||  |:| | ||||  :| || :: || || |||| |:|| |||||||
Db    849 PDPEVYQHEIPGGQLTNLLFQAQVGLGEKWVETKKAYFAANRLLGDIVKVTPTSKVVGD 908
Qy    903 LALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSE-- 959
           | || ||||  ||  ||  | |   :|  |||| || | |||| |||| ||| |
Db    909 LAQPMVSNKLSSEDVERLASELDFPDSVLDFFEGLMGTPYGGFPEPLRTNVISGKRRKLT 968
Qy    960 GKAPLTEVPEEEQA--HLDADDSKERRNSL-NRLLPKPTEEFLEHRRRFGNTSALDDR 1015
           |  | :| ||  :    |   |: : | :  ||: |  |||| : | | ::| |
Db    969 SRPGLTLEPYNIPAIREDLEARFSKVTENDVASYNMYPKVYEAYKKQQELYGDLSVLPTR 1028
Qy   1016 EFFY--GLVEGRETLIRLPDVRTPLLVRLDAISE-PDDKGMRNVVANVNGQIRPMRVRDR 1072
           | |    |:  || |:|| :||   |:|  |:   |   | :|: | ||| | | | |
Db   1029 NFLSPPKIDEERHVTVTIETRKTLIIKCMAEGELSQSSGTREVYFELNGEMRKVTVEDK 1088
Qy   1073 SVESVTATAEKADSSNKGHVAAPPAGVVT-VTVAEGDEVKAGDAVAIIEAMKMEATITAS 1131
           :  |  | :  || |||  |:   |||    | |  |:| |||| | |: |   ::|: |
Db   1089 NGAVETIIRPKADAHNPEIGAPMAGVVVEVRVHENGEVKKGDPIAVLSAMKMEMVISSP 1148
Qy   1132 VDGKIERVVVPAATKVEGGDLI 1153
           | | |:|   | |||  || ||
Db   1149 VAGRIGQIAVKENDSVDASDLI 1170
```

```
RESULT 7
PYCA_METJA
ID   PYCA_METJA     STANDARD;      PRT;    501 AA.
AC   Q58626;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pyruvate carboxylase subunit A (EC 6.4.1.1) (Pyruvic carboxylase A).
GN   PYCA OR MJ1229.
OS   Methanococcus jannaschii.
OC   Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC   Methanocaldococcaceae; Methanocaldococcus.
OX   NCBI_TaxID=2190;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX   MEDLINE=96337999; PubMed=8688087;
RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA   Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
```

```
Qy  208  ATEASREAEAAFGDGSVYYERAVINPQHIEVQILGDRIGEVHLYERDCSLQRRHQKVVE  267
         | ||||| :| ||||| || ||:|||||||| ||:|| |  ||:||:|||||||||:|
Db  181  IESARNIAKSAFGDPTVFIEKYLENPRHIEIQLLGDKHGNIIHLGDRECSIQRRHQKLIE  240

Qy  268  IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT  327
         || :  ||| :| ||: |: ||||| |||||||||| |||||||||||||:|||||||||
Db  241  EAPSPIMTEELRERMGEAAIKAGKAINYDSAGTVEFLY-ENGNFYFLEMNTRIQVEHTVT  299

Qy  328  EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA  387
         |:|| ||||||||:||||||:||| |:|  ::||| |:| ||| ||| |||||||
Db  300  EQVTGIDLVKAMIKIAAG---EELTLKQEDVKIRGHAIECRINAEDPLNDFVPCGKIKL  356

Qy  388  YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT  447
         |||||||||||  || ::|||| |:||| |||||||||||||||||||||||||||||
Db  357  YRSPGGPGVRIDSGVYGGAEIPPYYDSMIAKLITYGNSREEAIARMKRALREYVIIGVKT  416

Qy  448  NIGFLRALLREEDFTSKRIATGFIGDHPHLLQ  479
         || ||| :| ||:|| ||:| |:|| :|
Db  417  NIPTHRAVLEEENFLKGNISTHYVEQNMHKLR  448


RESULT 8
ACCC_ANASP
ID   ACCC_ANASP    STANDARD;    PRT;    447 AA.
AC   Q06862;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA
DE   carboxylase) (EC 6.4.1.2) (ACC).
GN   ACC OR ALR0939.
OS   Anabaena sp. (strain PCC 7120).
OC   Bacteria; Cyanobacteria: Nostocales; Nostocaceae; Nostoc.
OX   NCBI_TaxID=103690;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93352435; PubMed=8102363;
RA   Gornicki P., Scappino L.A., Haselkorn R.;
RT   "Genes for two subunits of acetyl coenzyme A carboxylase of Anabaena
RT   sp. strain PCC 7120: biotin carboxylase and biotin carboxyl carrier
RT   protein.";
RL   J. Bacteriol. 175:5268-5272(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21595285; PubMed=11759840;
RA   Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA   Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakazaki N., Shinpo S., Sugimoto M., Takazawa M., Yamada M.,
RA   Yasuda M., Tabata S.;
RT   "Complete genomic sequence of the filamentous nitrogen-fixing
RT   cyanobacterium Anabaena sp. strain PCC 7120.";
RL   DNA Res. 8:205-213(2001).
CC   -!- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
CC       CARBOXYLASE COMPLEX: FIRST, BIOTIN CARBOXYLASE CATALYZES THE
CC       CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE
CC       TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA.
CC   -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC       = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC   -!- PATHWAY: Long-chain fatty acid biosynthesis; first step.
CC   -!- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN
CC       CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
CC       OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX.
CC   -!- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
CC       PHOSPHATE SYNTHETASES.
```

```
DR   EMBL; L14862; AAB51770.1; -.
DR   EMBL; AP003584; BAB72896.1; -.
DR   PIR; A53311; A53311.
DR   PIR; AH1923; AH1923.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR004549; AccC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   TIGRFAMs; TIGR00514; accC; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Fatty acid biosynthesis; Ligase; Biotin; ATP-binding;
KW   Complete proteome.
FT   NP_BIND     163    168    ATP (BY SIMILARITY).
FT   ACT_SITE    293           BY SIMILARITY.
SQ   SEQUENCE   447 AA;   49104 MW;   8A541B3B39E00F9 CRC64;

Query Match            17.6%;   Score 1032.5;   DB 1;   Length 447;
Best Local Similarity  48.2%;   Pred. No. 2e-47;
Matches 216;   Conservative 71;   Mismatches 148;   Indels 13;   Gaps 4;

Qy   29  FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRICTEGSPVKAY   88
         ||||:|||||||||:|||:| | ||:| || |||||  ::| | ||||| |:|||||||
Db    3  FDKILIANRGEIALRILRACEEMGIATIAVHSTVDRNALHVQLADEAVCIG-EPASAKSY   61

Qy   89  LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV  148
         | ||| |||| :| ||||:|||||||||||:| || ||:|||||||||:|||||||:|||
Db   62  LNIPNIIAALTRNASAIHPGYGFLSENAKPAEICADHHIAFIGPTPEAIRLMGDKSTAK  121

Qy  149  TAAKKAGLPVLAES-----TPSKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDE  203
         : |  ||    ||:      :::   : | || :::|||||||||||||||||| ::
Db  122  ETMQKAGVPTVPGSEGLVETEQEGLE---LAKDIGYPVMIKATAGGGRGMRLVRSPDE  177

Qy  204  LRKLATEASREAEAAFGDGSVYYERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQ  263
         | ||| || |  |||||| ||||||||||||:|||||||| |:|||| || ||||||
Db  178  FVKFLAAQGEAGAAFGNAGVYIEKFIERPRHIEFQILADNYGNVIHLGERDCSIQRRNQ  237

Qy  264  KVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVE  323
         :||: ||:| |:| |||||||||||||:|| | |||||||||||:|||||| |:|||
Db  238  KLLEEAPSPALDSDLREKMGQAAVKAAQFINYTGAGTIEFLLDRSGQFYFEMEMNTRIQVE  297

Qy  324  HTVTEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTG  383
         | |||| ||||||: :| |||| ||| ||::| ||  || | |::| ||| ||
Db  298  HPVTEMVTGVDLLVEQIRIAQGERLR---LTQDQVVLRGHAIECRINAEDPHDFRPAPG  354

Qy  384  TITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVS  443
         :||:|||||||||||||||||||||||||:||||||||||||||||||||||||:
Db  355  RISGYLPPGGPGVRIDSHVTDYQIPPYYDSLIGKLIVWGPDRATAINRMKRALRECAIT  414

Qy  444  GVATNIGFLRALLREEDFTSKRIATGFI  471
         || || ||| :| ||:||:||:| |:
Db  415  GLPTTIGFHQRIMENPQFLQGNVSTSFV  442


RESULT 9
PYCA_ARCFU
ID   PYCA_ARCFU    STANDARD;    PRT;    506 AA.
AC   030019;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pyruvate carboxylase subunit A (EC 6.4.1.1) (Pyruvic carboxylase A).
GN   PYCA OR AF0220.
OS   Archaeoglobus fulgidus.
OC   Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC   Archaeoglobaceae; Archaeoglobus.
OX   NCBI_TaxID=2234;
```

```
Db  121 ESMQRAGVPVIPGSPKIDTVDEAKEWAEKIGYPVAVKASGGGGIGIVVNSQEELEEAF 180
Qy  209 TEASREAEAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEI 268
Db  181 RKSKKLGESYFKDSTVYLEKYLARPRHIEVQILADQHGNVIHLGERECSIQRRHQKLIEE 240
Qy  269 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE 328
Db  241 APSPALNEEMREELGKLAVKGAREIGYTNAGTFEFLY-ENGNFYLEINSRLQVEHTITE 299
Qy  329 EVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAY 388
Db  300 VVTGIDIVKYQIRIAYG---EELRHGQEDVAIRGHAIECRINAEDPVN-FYPRSGRILHY 355
Qy  389 RSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATN 448
Db  356 RSPGGIGIRVDSGIHMGYRIPEEYDSMISKLIAYGETREEAIARMKRALYEYIIEGVETN 415
Qy  449 IGFLRALLREEDFTSKRIATGFI 471
Db  416 IPFHFAVLNDEEFVRGNIHTKFV 438


RESULT 10
PYCA_METTH

ID   PYCA_METTH     STANDARD;      PRT;   491 AA.
AC   O27939;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pyruvate carboxylase subunit A (EC 6.4.1.1) (Pyruvic carboxylase A).
GN   PYCA OR MTH1917.
OS   Methanobacterium thermoautotrophicum.
OC   Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC   Methanobacteriaceae; Methanothermobacter.
OX   NCBI_TaxID=187420;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Delta H;
RX   MEDLINE=98037514; PubMed=9037463;
RA   Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA   Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA   Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA   Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA   Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA   McDougall S., Shiner G., Goyal A., Pietrovski S., Church G.M.,
RA   Daniels J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT   "Complete genome sequence of Methanobacterium thermoautotrophicum
RT   deltaH: functional analysis and comparative genomics.";
RL   J. Bacteriol. 179:7135-7155(1997).
RN   [2]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-17.
RC   STRAIN=Delta H;
RX   MEDLINE=98148063; PubMed=9478969;
RA   Mukhopadhyay B., Stoddard S.F., Wolfe R.S.;
RT   "Purification, regulation, and molecular and biochemical
RT   characterization of pyruvate carboxylase from Methanobacterium
RT   thermoautotrophicum strain deltaH.";
RL   J. Biol. Chem. 273:5155-5166(1998).
CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE CARBOXYL
CC       GROUP TO PYRUVATE IN THE SECOND. THE MAXIMUM ACTIVITY IS AT PH 8
CC       AND 60 DEGREES CELSIUS.
CC   -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
CC       oxaloacetate.
CC   -!- COFACTOR: ATP, MAGNESIUM (OR MANGANESE OR COBALT), PYRUVATE AND
CC       BICARBONATE.
CC   -!- ENZYME REGULATION: INHIBITED BY ADP AND ALPHA-KETOGLUTARATE.
CC   -!- PATHWAY: Gluconeogenesis.
CC   -!- SUBUNIT: HETEROOCTAMER OF FOUR A AND FOUR B SUBUNITS.
CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX   MEDLINE=98049343; PubMed=9389475;
RA   Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA   Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA   Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA   Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA   Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA   Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA   Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA   Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA   Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA   Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA   Venter J.C.;
RT   "The complete genome sequence of the hyperthermophilic, sulphate-
RT   reducing archaeon Archaeoglobus fulgidus.";
RL   Nature 390:364-370(1997).
CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE CARBOXYL
CC       GROUP TO PYRUVATE IN THE SECOND.
CC   -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
CC       oxaloacetate.
CC   -!- COFACTOR: ATP, MAGNESIUM (OR MANGANESE OR COBALT), PYRUVATE AND
CC       BICARBONATE (BY SIMILARITY).
CC   -!- PATHWAY: Gluconeogenesis.
CC   -!- SUBUNIT: HETEROOCTAMER OF FOUR A AND FOUR B SUBUNITS (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC       AND CARBAMYL PHOSPHATE SYNTHETASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AE001090; AAB91012.1; -.
DR   PIR; D69277; D69277.
DR   HSSP; P24182; 1BNC.
DR   TIGR; AF0220; .
DR   InterPro; IPR004549; AccC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR005479; Biotin_carb_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00289; Biotin_carb_D2; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   TIGRFAMs; TIGR00514; accC; 1.
DR   PROSITE; PS00866; CPSASE_1; FALSE_NEG.
DR   PROSITE; PS00867; CPSASE_2; FALSE_NEG.
KW   Ligase; Multifunctional enzyme; Gluconeogenesis; Magnesium; Pyruvate;
KW   ATP-binding; Complete proteome.
FT   NP_BIND     161    166       ATP (POTENTIAL).
FT   ACT_SITE    290    290       POTENTIAL.
SQ   SEQUENCE   506 AA;  57438 MW;  39B765F319235AD1 CRC64;

Query Match             17.4%;  Score 1020;  DB 1;  Length 506;
Best Local Similarity   47.9%;  Pred. No. 1.1e-46;
Matches  212;  Conservative   76;  Mismatches 149;  Indels   6;  Gaps   4;

Qy   29 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 88
Db    2 FSKILVANRGEIAVRVRVMRACRELGIKTVGVYSSADKRAFHRVYADECYIG-KADPRDSY 60
Qy   89 LDIDEIIGAAKVKADAIYPGVFGFLSENAQLARECAENGITFIGPTEVLDLTGDKSRAV 148
Db   61 LNIDRIIEVAKKSGAEAIHPGYGFLAENAEFAERCEEGIVFIGPSPEVIRIAGSKVRSR 120
Qy  149 TAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDELRKLA 208
```

```
CC          AND CARBAMYL PHOSPHATE SYNTHETASES.
CC   ------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use by  non-profit  institutions as long  as its content  is in  no  way
CC   modified and this statement is not removed.  Usage  by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------
DR   EMBL; AE000942; AAB86377.1; -.
DR   PIR; A69123; A69123.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR004549; AccC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   TIGRFAMs; TIGR00514; accC; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Ligase; Multifunctional enzyme; Gluconeogenesis; Magnesium; Pyruvate;
KW   ATP-binding; Complete proteome.
FT   NP_BIND     162    167       ATP (BY SIMILARITY).
FT   ACT_SITE    291    291       POTENTIAL.
SQ   SEQUENCE   491 AA;  54656 MW;  5789C34DA7475C2E CRC64;

Query Match             17.0%;  Score 996;  DB 1;  Length 491;
Best Local Similarity   45.6%;  Pred. No. 1.9e-45;
Matches  204;  Conservative  81;  Mismatches  156;  Indels   6;  Gaps   4;

Qy         29 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRGTEGSPVKAY 88
              |||||||||||| ::|:|  :  ||| : |:|  :||||: ||::: :|||::|::::|
Db          2 FSKILVANRGEIAIRVMRACRELGIKSVAVVSEADKNALFTRYADEAYEIG-KPAPSQSY 60

Qy         89 LDIDEIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 148
              | ||: |: ||||||||||||||||: :: ||| |:|||||||||||||||||||||||
Db         61 LRIDRILEVAEKAGAEAIIPGYGFLAENPRLGEECEKQGIKLIGPKGSVIEAMGDKITSK 120

Qy        149 TAAKKAGLPVLAESTPS-KNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDELRKL 207
              ||:|: |::|:  :: : :::|:  |:||: :||||:|||:|||:|||||: |||::|
Db        121 KLMKKAGVPVLPGTDQGVSDPDEARIADSIGYPVIIKASAGGGGIGMRAVYEEDELIRA 180

Qy        208 ATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVE 267
              |:: ||| || ||:::|:|:| :::| |:|:|:|:||  ||:::  |||:||::||::
Db        181 MESTQSVAASAFGDPTVYIEKYLERPRHIEFQVMADESGNVIHLADRECSIQRRHQKLIE 240

Qy        268 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 327
              |:   |:||  |::  ::| |:| ::|||:||||||||||:|| | |||||||||||:|
Db        241 EAPSPIMTPELRERMGSAAVKAAEYIGYENAGTVEFLY-SNGDFYFLEMNTRIQVEHPIT 299

Qy        328 EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 387
              ||:||||| |::::|: :|| : |: |:||:|:| ||:|  |:|||||::|||::||:||
Db        300 EVITGVDLVKEQIRVASG--EELRFTQKDINIRGHAIECRINAENPLADFAPNPGKITG 356

Qy        388 YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 447
              ||||||:||::||||||:|||||||||:||:|::|: ||::|:::||:|:||:||:||||
Db        357 YRSPGGIGVRVDSGVYMNYEIPPFYDSMISKLIVWGMDRQEAINRMKRALSEYIILGVKT 416

Qy        448 NIGFLRALLREEDFTSKRIATGFIGDH 474
              |||| || |:::|::| :  :: :: :
Db        417 TIPFHKAIMRNEAFRRGELHTHFVDEY 443
```

```
RESULT 11
ACCC_HAEIN
ID   ACCC_HAEIN     STANDARD;      PRT;   448 AA.
AC   P43873;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA
DE   carboxylase) (EC 6.4.1.2) (ACC).
GN   ACCC OR HI0972.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Rd / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT   Rd.";
RL   Science 269:496-512(1995).
CC   -!- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
CC       CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE
CC       CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE
CC       TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC       = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC   -!- PATHWAY: Long-chain fatty acid biosynthesis; first step.
CC   -!- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN
CC       CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
CC       OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
CC       PHOSPHATE SYNTHETASES.
CC   ------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is  produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use by  non-profit  institutions as long  as its content  is in  no  way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------
DR   EMBL; U32778; AAC22632.1; -.
DR   PIR; F64105; F64105.
DR   HSSP; P24182; 1BNC.
DR   TIGR; HI0972;
DR   InterPro; IPR004549; AccC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   TIGRFAMs; TIGR00514; accC; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Fatty acid biosynthesis; Ligase; Biotin; ATP-binding;
KW   Complete proteome.
FT   NP_BIND     163    168       ATP (POTENTIAL).
FT   ACT_SITE    292    292       POTENTIAL.
SQ   SEQUENCE   448 AA;  49108 MW;  2B497E2A31ED96D1 CRC64;

Query Match             16.2%;  Score 950;  DB 1;  Length 448;
Best Local Similarity   48.3%;  Pred. No. 4.5e-43;
Matches  212;  Conservative  57;  Mismatches  152;  Indels  18;  Gaps   9;

Qy         30 KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRGTEGSPVKAYL 89
              :::|||||||||||:| | | | | ||:: | | || | : |  :  |:||
Db          3 EKVVIANRGEIAILRILRACKELGIKTVAVHSTADRDLKHVLLADETICIGPAPS-AKSYL 61

Qy         90 DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT 149
```

```
Db           :  ||| :  |||:  ||||:||||||||  |:   : |||| :
          62 NIPAIIAAAEVTGADAIHPGYGFLSENADFAEQVERSGFTFIGPTADVIRLMGDKVSAIK 121

Qy   150 AAKKAGLPVLAEST-PSKNIDDIVKS--AEGQTYPIFVKAVAGGGRGMRFVSSPDELR 205
          |  ||||:  :|  |: |||||:  :||  ||||||||||  | |
Db   122 AMKKAGVPCVPGSDGPVSN-DIAKNKEIAKRIGYPIIIKASGGGGRGMRVVRSEDALE 179

Qy   206 KLATEASREAEAAFGDGSVVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKV 265
          ||:  |  |||:  :  |||:  : |:||:::|| |  |||||||||||
Db   180 ESIAMTKAFAKAAFNDMVYMEKYLENPRHVEIQVLADTHGNAVYLAERDCSMQRRHQKV 239

Qy   266 VEIAPAQHLDPELRDRI---CADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQV 322
          ||  : :|  : |:|:  || |  || |||| | || |||| ||| |||||||
Db   240 VEEAPAPGITEEVRRDIGSRCANA---CVEIGYRGAGTEEFLY-ENGEFYFIEMNTRIQV 295

Qy   323 EHTVEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDT 382
          ||  |: || :| :||||  | |||||:||  || :|  |: |  :|||:|
Db   296 EHPVTEMITGVDLVKEQLRIAAGL---PISFKQEDIKVKGHAMECRINAEDPKT-FLPSP 351

Qy   383 GTIIAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTV 442
          |  :|| || ||||| ||  |:|  | :|| |  ||||:  |  :| |:  |:|
Db   352 GKVNHLHSPGGLGVRWDSHYYGGYTVPPHYDSMIAKLITYGDTREVAIRRMQNALSETII 411

Qy   443 SGVATNIGFLRALLREEDF 461
          |||  ||| |||||::||
Db   412 DGIKTNIPLHELLLEDENF 430


RESULT 12
ACCC_ECOLI
ID   ACCC_ECOLI     STANDARD;      PRT;    449 AA.
AC   P24182;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA
DE   carboxylase) (EC 6.4.1.2) (ACC).
GN   ACC OR FABG OR B3256.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=K12;
RX   MEDLINE=92052166; PubMed=1682920;
RA   Tsuru D., Anai M., Sekiguchi M., Yoshimoto T., Masuda T., Kitazono A.,
RA   Kondo H., Shiratsuchi K., Yoshimoto T., Masuda T., Tanabe T.;
RT   "Acetyl-CoA carboxylase from Escherichia coli: gene organization and
RT   nucleotide sequence of the biotin carboxylase subunit.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:9730-9733(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92112819; PubMed=1370469;
RA   Li S.-J., Cronan J.E. Jr.;
RT   "The gene encoding the biotin carboxylase subunit of Escherichia coli
RT   acetyl-CoA carboxylase.";
RL   J. Biol. Chem. 267:855-863(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Best E.A., Knauf V.C.;
RL   Submitted (XXX-1992) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=94726617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [5]
RP   SEQUENCE OF 1-12.
RC   STRAIN=K12 / EMG2;
RX   MEDLINE=97443975; PubMed=9298646;
RA   Link A.J., Robison K., Church G.M.;
RT   "Comparing the predicted and observed properties of proteins encoded
RT   in the genome of Escherichia coli K-12.";
RL   Electrophoresis 18:1259-1313(1997).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX   MEDLINE=94347755; PubMed=7915138;
RA   Waldrop G.L., Rayment I., Holden H.M.;
RT   "Three-dimensional structure of the biotin carboxylase subunit of
RT   acetyl-CoA carboxylase.";
RL   Biochemistry 33:10249-10256(1994).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX   MEDLINE=20283656; PubMed=10821865;
RA   Thoden J.B., Blanchard C.Z., Holden H.M., Waldrop G.L.;
RT   "Movement of the biotin carboxylase B-domain as a result of ATP
RT   binding.";
RL   J. Biol. Chem. 275:16183-16190(2000).
CC   -!- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
CC       CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE
CC       CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE
CC       TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA.
CC   -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC       = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC   -!- PATHWAY: Long-chain fatty acid biosynthesis; first step.
CC   -!- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN
CC       CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
CC       OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX.
CC   -!- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
CC       PHOSPHATE SYNTHETASES.
CC   -------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------------
DR   EMBL; M79446; AAA23748.1; -.
DR   EMBL; M80458; AAA23409.1; -.
DR   EMBL; M83198; AAA23746.1; -.
DR   EMBL; U18997; AAA58059.1; -.
DR   EMBL; AE000404; AAC76288.1; -.
DR   PIR; JS0632; JS0632.
DR   PDB; 1BNC; 30-AUG-95.
DR   PDB; 1DV1; 09-JUN-00.
DR   PDB; 1DV2; 09-JUN-00.
DR   PDB; 1K69; 05-DEC-01.
DR   EcoGene; EG10276; accC.
DR   InterPro; IPR004549; AccC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR005479; Biotin_carb_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   TIGRFAMs; TIGR00514; accC; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Fatty acid biosynthesis; Ligase; Biotin; ATP-binding; 3D-structure;
KW   Complete proteome.
FT   NP_BIND     163    168       ATP (POTENTIAL).
FT   ACT_SITE    292    292       POTENTIAL.
FT   CONFLICT    260    261       CA -> SR (IN REF. 2).
FT   CONFLICT    313    313       L -> M (IN REF. 1).
FT   STRAND        4      7
FT   HELIX        11     24
FT   TURN         25     25
FT   STRAND       27     33
```

```
FT   HELIX     34     36
FT   TURN      37     38
FT   HELIX     40     44
FT   STRAND    47     52
FT   HELIX     56     58
FT   TURN      59     61
FT   HELIX     63     73
FT   TURN      74     74
FT   STRAND    77     79
FT   TURN      84     87
FT   HELIX     89     97
FT   TURN      98     99
FT   HELIX     99    102
FT   STRAND   101    114
FT   HELIX    107    114
FT   HELIX    116    125
FT   TURN     126    127
FT   STRAND   130    130
FT   STRAND   133    133
FT   HELIX    142    152
FT   STRAND   156    160
FT   TURN     172    172
FT   HELIX    175    188
FT   STRAND   200    202
FT   STRAND   208    217
FT   STRAND   222    234
FT   TURN     235    236
FT   STRAND   237    243
FT   TURN     246    247
FT   TURN     250    267
FT   TURN     268    268
FT   STRAND   270    280
FT   STRAND   281    282
FT   STRAND   283    290
FT   TURN     295    296
FT   HELIX    297    304
FT   HELIX    308    316
FT   TURN     317    318
FT   HELIX    325    327
FT   STRAND   333    340
FT   STRAND   342    346
FT   TURN     344    346
FT   STRAND   349    349
FT   STRAND   352    353
FT   STRAND   356    358
FT   TURN     363    364
FT   STRAND   365    368
FT   TURN     373    374
FT   STRAND   376    377
FT   STRAND   384    392
FT   HELIX    395    408
FT   STRAND   410    412
FT   HELIX    418    425
FT   TURN     426    426
FT   HELIX    428    432
FT   HELIX    437    438
FT   HELIX    439    444
FT   TURN     445    445
SQ   SEQUENCE   449 AA;   49320 MW;  68C55F10ACB4F170 CRC64;

Query Match             16.0%;     Score 940;   DB 1;   Length 449;
Best Local Similarity   46.8%;     Pred. No. 1.5e-42;
Matches  203;   Conservative   66;   Mismatches  155;   Indels  10;   Gaps   6;

Qy      31  KILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLD   90
            ||::||||:|| :               ||| | |||||||| | ||||
Db       4  KIVIANRGEIALRILRACKELGIKTVAVHSSADRDLKHVLLADETVCIGPAPS-VKSYLN   62

Qy      91  IDEIIGAAKVKADAIYPGYGVGFLSENAQLARECAENGITFIGPTPEVLLTGDKSRAVTA  150
            | ||  |||||||||||||||||||| | |||   |||||||| |||||| | |||| |
Db      63  IPAIISAAEITGAVAIHPGYGFLSENANFAEQVERSGFIFIGPKAETIRLMGDKVSAIAA  122

Qy     151  AKKAGLPLAESTPSKNIDDIVKS--AEGQTYPIFVKAVAGGGRGMRFVSSPDELRKL  207
            |||||:| :  ||  || : || ::   :  ||  |||||||||||||| | ||
```

```
Db     123  MKKAGVPCVPGSDGPLG-DDMDKNRAIAKRIGYPVIIKASGGGGRGMRVVRGDAELAQS  181
            |||||| | |   | |   || ||||||: | |||||||:|||||||||:||  ||| |
Qy     208  ATEASREAEAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVE  267
            |   |::         |||||||  ||:|:::||||||| :|:| |||:||:| |:||:
Db     182  ISMTRAEAKAAFSNDMVYMEKYLENPRHVEIQVLADGQGNAIYLAERDCSMQRRHQKVVE  241
            || :||| | |:|| |: |:||||| | |:||:|| | | | ||:|| || :|| ||||
Qy     268  IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT  327
            ||| | |  || | |   |  ||| || ||| ||| | | | || || | ||||| ||
Db     242  EAPAPGITPELRRYIGERCAKACVDIGYRGAGTFEFLF-ENGEFYFIEMNTRIQVEHPVT  300
            ||||| || ||||  |||| || ||:||| |||| ||: || |:|||||| |||||: ||
Qy     328  EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGITA  387
            || ||||| |||  || |  |:||: |  |  |||    | | ||| |  |  | |  |
Db     301  EMITGVDLIKEQLRIAAG---QPLSIKQEEVHVRGHAVECRINAEDPNT-FLPSPGKITR  356
            |: ||||: |:|: :|||    | |    ||  | |  |||| |||| :  | |  | |
Qy     388  YRSPGGAGVRLDGAAQLGGETAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT  447
            ::| || ||    | ::   | :|||: : : ||| ||:|   | :|| |:: | ||:|
Db     357  FHAPGGFGVRWESHIYAGYTVPPYYDSMIGKLICYGENRDVAIARMKNALQELIIDGIKT  416
            | |||| ||| |: : |||:||| |||:|: |:| ||:| |:|||:||||:|:||||:|
Qy     448  NIGFLRALLREDF  461
            || :| | :| |
Db     417  NVDLQIRIMNDENF  430
```

```
RESULT 13
ACCC_ECO57
ID   ACCC_ECO57            STANDARD;      PRT;       449 AA.
AC   Q8X9B6;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA
DE   carboxylase) (EC 6.4.1.2) (ACC).
GN   ACCC OR 24616 OR ECS4128.
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   Nature 409:529-533(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
CC   -!- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
CC       CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE
CC       CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE
CC       TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA.
CC   -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC       = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC   -!- PATHWAY: Long-chain fatty acid biosynthesis; first step.
CC   -!- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN
CC       CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
CC       OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX (By similarity).
CC   -!- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
CC       PHOSPHATE SYNTHETASES.
------------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AE005553; AAG58384.1; -.
DR   EMBL; AP002564; BAB37551.1; -.
DR   PIR; H91144; H91144.
DR   InterPro; IPR004549; AccC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   TIGRFAMs; TIGR00514; accC; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Fatty acid biosynthesis; Ligase; Biotin; ATP-binding;
KW   Complete proteome.
FT   NP_BIND       163    168       ATP (POTENTIAL).
FT   ACT_SITE      292    292       POTENTIAL.
SQ   SEQUENCE   449 AA;  49337 MW;  BC5716323F6233ED CRC64;

Query Match             16.0%;  Score 938;  DB 1;  Length 449;
Best Local Similarity   46.8%;  Pred. No. 1.9e-42;
Matches  203;  Conservative  66;  Mismatches 155;  Indels 10;  Gaps  6;

Qy          31  KILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLD   90
                ||:||||||||   ||: : |||||| ||  :||:||  |:|| |:||||||||||||||
Db           4  KIVIANRGEIALRILRACKELGIKTVAVHSSADRDLKHVLLADETVCIGPAPS-VKSYLN   62

Qy          91  IDEIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLLTGDKSRAVTA   150
                ||::| |  |||| |||:||||||||||  || :| |||| | | |:|||||||:|| |
Db          63  IPAIISAAEITGAVAIHPGYGFLSENANFAEQVERSGFIFIGPKAETIRLMGDKVSAIAA   122

Qy         151  AKKAGLPVLAESTPSKNIDDIVKS---AEGQTYPIFVKAVAGGGRGMRFVSSPDELRKL   207
                //||:: : ||:  || | :|: ||  :|| :| ||||| | :|||||||  ||| |:
Db         123  MKKAGVPCVPGSDGPLG-DDMDKNRAIAKRIGYPVIIKASGGGGRGMRVVRGDAELAQS   181

Qy         208  ATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVE   267
                |||:||||:|||||||||| || |:||: |||:||:||| |||||:||| |:||||||||
Db         182  ISMTRAEAKAAFSNDMVYMEKYLENPRHVEIQVLADGQGNSIYLAERDCSMQRRHQKVVE   241

Qy         268  IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT   327
                |||  : |||| |  ||:: |||||||||||||||:|:||:|||||| |||||||||||
Db         242  EAPAPGITPELRRYIGERCAKACVDIGYRGAGTFEFLF-ENGEFYFIEMNTRIQVEHPVT   300

Qy         328  EEVTEVDLVKAQMRLAAGATLKEIGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA   387
                |||:| :|||:|||   ||| ::|||:  :: | | |:|  |||:||||||:| ||| ||
Db         301  EMITGVDLIKEQLRIAAG---QPLSIKQEFVHVRGHAVECRINAEDPNT-FLPSPGKITR   356

Qy         388  YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT   447
                ||||||||| ||| || ||||| ||||  | ||: | |:| ||::||| ||||||||||:
Db         357  FHAPGGFGVRWESHIYAGYTVPPYYDSMIGKLICYGENRDVAIARMKNALQELIIDGIKT   416

Qy         448  NIGFLRALLREEDE   461
                |: :| :|: :|
Db         417  NVDLQIRIMNDENF   430

RESULT 14
ACCC_PSEAE
ID   ACCC_PSEAE     STANDARD;      PRT;   449 AA.
AC   P37798;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA
DE   carboxylase) (EC 6.4.1.2) (ACC).
GN   ACCC OR FABG OR PA4848.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=94042851; PubMed=7693652;
RA   Best E.A., Knauf V.C.;
RT   "Organization and nucleotide sequences of the genes encoding the
RT   biotin carboxyl carrier protein and biotin carboxylase protein of
RT   Pseudomonas aeruginosa acetyl coenzyme A carboxylase.";
RL   J. Bacteriol. 175:6881-6889(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L.N., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
-!- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
CC       CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE
CC       CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE
CC       TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA.
-!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC       = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
-!- PATHWAY: Long-chain fatty acid biosynthesis; first step.
-!- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN
CC       CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
CC       OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX.
-!- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
CC       PHOSPHATE SYNTHETASES.
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch).
CC   or send an email to license@isb-sib.ch).
DR   EMBL; L14612; AAA16041.1; -.
DR   EMBL; AE004898; AAG08233.1; -.
DR   PIR; B49342; B49342.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR004549; AccC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   TIGRFAMs; TIGR00514; accC; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Fatty acid biosynthesis; Ligase; Biotin; ATP-binding;
KW   Complete proteome.
FT   NP_BIND       163    168       ATP (POTENTIAL).
FT   ACT_SITE      292    292       POTENTIAL.
SQ   SEQUENCE   449 AA;  48888 MW;  3B04C77785C73541 CRC64;

Query Match             15.7%;  Score 921;  DB 1;  Length 449;
Best Local Similarity   45.2%;  Pred. No. 1.5e-41;
Matches  196;  Conservative  72;  Mismatches 158;  Indels  8;  Gaps  5;

Qy          30  KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSPASEAVRIGTEGSPVKAYL   89
```

```
RESULT 15
MCCA_HUMAN
ID   MCCA_HUMAN     STANDARD;      PRT;   725 AA.
AC   Q96RQ3; Q9H959; Q9NS97;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Methylcrotonoyl-CoA carboxylase alpha chain, mitochondrial precursor
DE   (EC 6.4.1.4) (3-Methylcrotonyl-CoA carboxylase 1) (MCCase alpha
DE   subunit) (3-methylcrotonyl-CoA:carbon dioxide ligase alpha subunit).
GN   MCCC1 OR MCCA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND VARIANTS MCGI ARG-325 AND SER-385.
RX   PubMed=1170888;
RA   Gallardo M.E., Desviat L.R., Rodriguez J.M., Esparza-Gordillo J.,
RA   Perez-Cerda C., Perez B., Rodriguez-Pombo P., Criado O., Sanz R.,
RA   Morton D.H., Gibson K.M., Le T.P., Ribes A., Rodriguez de Cordoba S.,
RA   Ugarte M., Penalva M.A.;
RT   "The molecular basis of 3-methylcrotonyl-CoA carboxylase deficiency:
RT   gene (MCCA): cDNA sequence, genomic organization, localization to
RT   chromosomal band 3q27, and expression.";
RL   Am. J. Hum. Genet. 68:334-346(2001).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT HIS-464.
RX   MEDLINE=21295033; PubMed=11401427;
RA   Obata K., Fukuda T., Morishita R., Abe S., Asakawa S., Yamaguchi S.,
RA   Yoshino M., Ihara K., Murayama K., Shigemoto K., Shimizu N., Kondo I.;
RT   "Human biotin-containing subunit of 3-methylcrotonyl-CoA carboxylase
RT   gene (MCCA): cDNA sequence, genomic organization, localization to
RT   leucine catabolism.";
RL   Genomics 72:145-152(2001).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANT MCGI PHE-535.
RX   MEDLINE=21299419; PubMed=11406611;
RA   Holzinger A., Roeschinger W., Lagler F., Mayerhofer P.U., Lichtner P.,
RA   Kattenfeld T., Thuy L.P., Nyhan W.L., Koch H.G., Muntau A.C.,
RA   Roscher A.A.;
RT   "Cloning of the human MCCA and MCCB genes and mutations therein reveal
RT   the molecular cause of 3-methylcrotonyl-CoA: carboxylase
RT   deficiency.";
RL   Hum. Mol. Genet. 10:1299-1306(2001).
RN   [4]
RP   SEQUENCE FROM N.A., AND VARIANTS MCGI VAL-289; SER-385; PRO-437 AND
RP   HIS-532.
RX   PubMed=1181649;
RA   Baumgartner M.R., Almashanu S., Suormala T., Obie C., Cole R.N.,
RA   Packman S., Baumgartner E.R., Valle D.;
RT   "The molecular basis of human 3-methylcrotonyl-CoA carboxylase
RT   deficiency.";
RL   J. Clin. Invest. 107:495-504(2001).
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA   Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA   Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skeletal muscle;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- CATALYTIC ACTIVITY: ATP + 3-methylcrotonyl-CoA + HCO(3)(-) = ADP +
CC       phosphate + 3-methylglutaconyl-CoA.
CC   -!- COFACTOR: Biotin.
CC   -!- PATHWAY: Leucine catabolism.
CC   -!- SUBUNIT: Probably a dodecamer composed of six biotin-containing
CC       alpha subunits and six beta subunits.
CC   -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC   -!- DISEASE: Defects in MCCC1 are the cause of 3-
CC       methylcrotonylglycinuria type I (MCGI) [MIM:210200]; also
CC       designated CGA or CG2. MCGI is a recessive disease that is
CC       characterized by muscular hypotonia and atrophy, probably of
CC       spinal origin.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF310972; AAG53095.1; -.
DR   EMBL; AB029826; BAA99407.1; -.
DR   EMBL; AF297332; AAK67986.1; -.
DR   EMBL; AF310339; AAG50245.1; -.
DR   EMBL; AK023051; BAB14377.1; -.
DR   EMBL; BC004214; AAH04214.1; -.
DR   EMBL; BC004187; AAH04187.1; -.
DR   Genew; HGNC:6936; MCCC1.
```

```
Qy    3 EKVLIANRGEIALRILRACKELGIKTVAVHSTADRELMHLSLADESVCIG-PAPATQSYL 61
          ::|:||||||| ||| ||  || ||  |:| ||
Db   90 DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDTGDKSRAVT 149

Qy   62 QIPAIIAAEVTGATAIHPGYGFLAENADFAEQIERSGFTFVGPTAEVIRLMGDKVSAKD 121
          |   |:||| :|||||||||||| :|::||||| |||| ||  ||||
Db  150 AAKKAGLPVL--AESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDELRKL 207

Qy  122 AMKRAGVPTVPGSDGPLPEDEETALAIAREVGYPVIIKAAGGGGRGMRVVYDESELIKS 181
          || | : ||||||||| :|::||||| |||| || ||||
Db  208 ATEASREEAAFGDGSVVVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVE 267

Qy  182 AKLRTEAGAAFGNPMVYLEKFLTNPRHVEVQVLSDGQGNAIIHLGRDCSLQRRHQKVIE 241
          | ::  : ||||||||:::||||| ||||||| :|::||||||
Db  268 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGKNHVFIEMNPRIQVEHTVT 327

Qy  242 EAPAPGIDEKARQEVFEARCVQACIEIGYRGAGTEFLY-ENGRYFIEMNTRVQVEHPVS 300
          | ||:| :| ||||| || |||| ||  ||||
Db  328 EEVEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 387

Qy  301 EMVTGVDIVKEMLRIASG---EKLSIRQEDVVIRGHALECRINAEDPKT-FMPSPGKVKH 356
          || | : |||||| |||| ||||||:::|
Db  388 YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 447

Qy  357 FHAPGGNGVRVDSHLYSGYSVPPNYDSLVGKVTTYGADRDEALARMRNALDELIVDGIKT 416
          | |:|: :|||||||:::|||
Db  448 NIGFLRALLREEDF 461

Qy  417 NTELHKDLVRDAAF 430
```

```
Qy   380 PDIGTITAYRSP-GGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRALA 438
             :: .: .:   :   ::.::: :.:.  :: : ::. :  ::    ::: .:
Db   397 PVAGPLVHLSTPRADPSTRIETGVRQGDEVSVHYDPMIAKLVVWAADRQAALTKLRYSLR 456

Qy   439 EFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVT 498
             :: .: ||| |||  :   : ||: .:     :  | :|.|
Db   457 QYNIVGLPTNIDFLLNLSGHPEFEAGNVHTDFIPQH------------------------ 492

Qy   499 VNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTTF 558
                            : :.:  :: | .:: |  |:||    :::  :|:
Db   493 ----H--------KQLLLSRKAAAK-ESLCQAALGLILKEK--AMTDFFT         527

Qy   559 RDAH---------QSLLATRVRSFALKPAAEAVA 583
         ||      :  :: ::  |  ||
Db   528 LQAHDQFSPFSSSSGRRLNISYTRNMTLKDGKNNVA 563
```

Search completed: September 24, 2003, 15:56:03
Job time : 22.6369 secs

```
DR   GK; Q96RQ3; -.
DR   MIM; 210200; -.
DR   GO; GO:0005759; C:mitochondrial matrix; NAS.
DR   GO; GO:0009374; F:biotin binding activity; NAS.
DR   GO; GO:0004485; F:methylcrotonoyl-CoA carboxylase activity; NAS.
DR   GO; GO:0006768; P:biotin metabolism; NAS.
DR   GO; GO:0006552; P:leucine catabolism; NAS.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Mitochondrion; Ligase; Biotin; ATP-binding; Transit peptide;
KW   Disease mutation; Polymorphism.
FT   TRANSIT       1     47       MITOCHONDRION (POTENTIAL).
FT   CHAIN        48    725       METHYLCROTONOYL-COA CARBOXYLASE ALPHA
FT                                CHAIN.
FT   NP_BIND     209    214       ATP (POTENTIAL).
FT   ACT_SITE    339    339       BY SIMILARITY.
FT   BINDING     681    681       BIOTIN (BY SIMILARITY).
FT   DOMAIN      538    541       POLY-SER.
FT   DOMAIN      713    718       POLY-GLU.
FT   VARIANT     289    289       A -> V (in MCGI; mild form).
FT                                /FTId=VAR_012785.
FT   VARIANT     325    325       M -> R (in MCGI).
FT                                /FTId=VAR_012786.
FT   VARIANT     385    385       R -> S (in MCGI; severe form).
FT                                /FTId=VAR_012787.
FT   VARIANT     437    437       L -> P (in MCGI; severe form).
FT                                /FTId=VAR_012788.
FT   VARIANT     464    464       P -> H.
FT                                /FTId=VAR_012789.
FT   VARIANT     532    532       D -> H (in MCGI; severe form).
FT                                /FTId=VAR_012790.
FT   VARIANT     535    535       S -> F (in MCGI; asymptomatic form).
FT                                /FTId=VAR_012791.
FT   CONFLICT    469    469       F -> L (IN REF. 3 AND 4).
SQ   SEQUENCE   725 AA;  80433 MW;  B847C6B8060686C0 CRC64;
```

Query Match   15.3%; Score 897; DB 1; Length 725;
Best Local Similarity   36.1%; Pred. No. 5.4e-40;
Matches 208; Conservative   96; Mismatches 206; Indels   66; Gaps   9;

```
Qy    21 HTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGT 80
            :::| |  |:||||||||:| :  | : | ||  | |: ::|| :     ::| ::  |
Db    41 YTTATGRNITKVLIANRGEIACRVMRTAKKLGVQTVAVYSEADRNSMHVDMADEAYSIGP 100

Qy    81 EGSPVKAYLDIDEITGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDL 140
         : :||  : :::|| ||||.||| ||||:|||||||| ::| ||:|:||:|| :|:| |
Db   101 APSQ-QSYLSMEKIIQVAKTSAAQAIHPGGGFLSENMEFAELCKQEGIIFIGPPPSAIRD 159

Qy   141 TGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKS-AEGQTYPIFVKAVAGGGGRGMRFVS 199
              :.|:   || ||:|| :   | ||:|::.  :|| || |:|||| |||||.|||:|
Db   160 MGIKSTSKSIMAAAGYPVVEGYHGEDQSDQCLKEHARRIGYPVMIKAVRGGGGKGMRIVR 219

Qy   200 SPDELRKLATEASREAEAAFGDGSVVVERAVINPQHIEVQILGDRTGEVHLYERDCSLQ 259
            :|  ::|| |:|:||   |:|.:: |:|: :|| | ||||::||: || :|||||::|
Db   220 SEQEFQEQLESARREAKKSFNDDAMLIEKFVDTPRHVEVQVFGDHHGNAVYLFERDCSVQ 279

Qy   260 RRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPR 319
         ||||::| |||| :: .:|:  : | |. :: :  |||||||:::|:|:|:|.|:|||:|
Db   280 RRHQKIIEEAPAPGIKSEVRKKLGEAAVRAAKAVNVVGAGTVEFIMDSKHNFCFMEMNTR 339

Qy   320 IQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFR 379
         :|||| ||| :|| |||  |:|:|||   : :|:|::  :| |||:| |||:|||.|.|:
Db   340 LQVEHPVTEMITGTDLVEWQLRIAAG---EKIPLSQEEITLQGHAFEARIYAEDPSNNFM 396
```

OM protein - protein search, using sw model

```
            GenCore version 5.1.6
       Copyright (c) 1993 - 2003 Compugen Ltd.
```

Run on: September 24, 2003, 15:51:28 ; Search time 89.6587 Seconds
(without alignments)
3330.038 Million cell updates/sec

Title: US-09-974-973A-2
Perfect score: 5865
Sequence: 1 MTAITLGGLLKGIITLVST.........RVVVPAATKVEGGDLIVVVS 1157

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
```
 1: sp_archea:*
 2: sp_bacteria:*
 3: sp_fungi:*
 4: sp_human:*
 5: sp_invertebrate:*
 6: sp_mammal:*
 7: sp_mhc:*
 8: sp_organelle:*
 9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	5772	98.4	1140	2	Q8L2G4	Q8l2g4 corynebacte
2	5759	98.2	1140	16	O54587	O54587 corynebacte
3	5322	90.7	1168	16	Q8FRQ0	Q8frq0 corynebacte
4	5297.5	90.3	1139	2	Q8RQL2	Q8rql2 corynebacte
5	3797	64.7	1124	16	Q9RK64	Q9rk64 streptomyce
6	3786.5	64.6	1127	2	Q9F843	Q9f843 mycobacteri
7	3713.5	63.3	1127	16	P95127	P95127 mycobacteri
8	3484	59.4	1124	2	Q50450	Q50450 mycobacteri
9	2550.5	43.5	1178	13	Q8JHF6	Q8jhf6 gallus gall
10	2535.5	43.2	1150	16	Q9K9M0	Q9k9m0 bacillus ha
11	2530	43.1	1148	16	Q9KWU4	Q9kwu4 bacillus su
12	2523.5	43.0	1180	13	Q9DDT1	Q9ddt1 brachydanio
13	2509	42.8	1147	2	P94448	P94448 bacillus st
14	2507.5	42.8	1192	3	Q9HES8	Q9hes8 aspergillus
15	2494	42.5	1146	16	Q8Y846	Q8y846 listeria mo
16	2491	42.5	1146	16	Q92CW1	Q92cw1 listeria in
17	2490.5	42.5	1144	16	Q97FR7	Q97fr7 clostridium
18	2490	42.5	1150	16	Q8NX69	Q8nx69 staphylococ
19	2489	42.4	1150	16	Q99UY8	Q99uy8 staphylococ
20	2485.5	42.4	1193	3	O93918	O93918 aspergillus
21	2478.5	42.3	1158	16	Q8FYT0	Q8fyt0 brucella su
22	2474.5	42.2	1174	16	Q8UBX3	Q8ubx3 agrobacteri
23	2473.5	42.2	1152	16	Q98F27	Q98f27 rhizobium l
24	2468.5	42.1	1179	13	Q8AYN3	Q8ayn3 pagrus majo
25	2465.5	42.0	1158	16	Q8YJ20	Q8yj20 brucella me
26	2465	42.0	1153	16	Q8CPM3	Q8cpm3 staphylococ
27	2461	42.0	1175	5	O17732	O17732 caenorhabdi
28	2447	41.7	1185	5	Q9UUE1	Q9uuel schizosacch
29	2441	41.6	1181	5	Q9XZ00	Q9xz00 drosophila
30	2439	41.6	1185	3	P78822	P78822 schizosacch
31	2435.5	41.5	1175	3	Q8X1T3	Q8x1t3 pichia angu
32	2435	41.5	1152	16	Q92L13	Q92l13 rhizobium m
33	2435	41.5	1197	5	Q8MKW5	Q8mkw5 drosophila
34	2417.5	41.2	1195	5	Q16921	Q16921 aedes aegyp
35	2417	41.2	1147	16	Q8ER83	Q8er83 oceanobacil
36	2413	41.1	1137	16	Q9CHQ7	Q9chq7 lactococcus
37	2405	41.0	1137	2	Q9RAT6	Q9rat6 lactococcus
38	2401	40.9	1154	2	Q59740	Q59740 rhizobium e
39	2155.5	36.8	1029	13	Q8AXQ6	Q8axq6 pagrus majo
40	2119.5	36.1	984	2	Q9XBJ1	Q9xbj1 bacillus ce
41	2021	34.5	935	11	Q62043	Q62043 mus musculu
42	1933.5	33.0	920	2	Q9KWU5	Q9kwu5 bacillus su
43	1661	28.3	828	11	Q8BP54	Q8bp54 mus musculu
44	1097	18.7	525	5	Q8MKW4	Q8mkw4 drosophila
45	1096	18.7	541	5	Q8MKW3	Q8mkw3 drosophila

ALIGNMENTS

RESULT 1
Q8L2G4
```
ID   Q8L2G4          PRELIMINARY;     PRT;   1140 AA.
AC   Q8L2G4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase.
GN   PYC.
OS   Corynebacterium crenatum.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=168810;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CD945;
RA   Wang J., Ding J., Liu Y.;
RT   "Cloning and Expression of Pyruvate Carboxylase Gene in
RT   Corynebacterium crenatum CD945.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF503915; AAM27458.1; -.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
SQ   SEQUENCE 1140 AA; 123126 MW; FFA90BB7644C910E CRC64;
```

```
Query Match          98.4%;    Score 5772;   DB 2;    Length 1140;
Best Local Similarity 99.6%;   Pred. No. 3.4e-300;
Matches 1135;   Conservative    4;   Mismatches   1;   Indels    0;   Gaps    0;

Qy    18 VSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 77
         ::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60

Qy    78 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 137
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120

Qy   138 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRF 197
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRF 180

Qy   198 VSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCS 257
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 VSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCS 240

Qy   258 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 317
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 300

Qy   318 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 377
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360

Qy   378 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 437
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420

Qy   438 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADV 497
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADV 480

Qy   498 TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 557
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 540

Qy   558 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 617
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 FRDAHQSLLATRVRSFALKPAAFAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600

Qy   618 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 677
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAATSGVDIFRIFDALNDVSQMRPA 660

Qy   678 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR 737
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR 720

Qy   738 PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 797
         |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 PAAATKLVTALRRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 780

Qy   798 LSAIVAAFAHTRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 857
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 LSAIVAAFAHTRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840

Qy   858 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPDTKVTPSSKVVGDLALHLVGAGVDPADF 917
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900

Qy   918 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA 977
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA 960

Qy   978 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1037
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020
```

```
Qy  1038 LLVRLDAISEPDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1097
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 LLVRLDAISEPDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080

Qy  1098 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATTASVDGK IERVVPAATKVEEGDLIVVVS 1157
         ||||||||||||||||||||||||||||||||||||:|||:::||||||||||||||:::
Db  1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATTASVDGKIDRVVPAATKVEEGDLIVVIS 1140


RESULT 2
054587
ID   054587;      PRELIMINARY;      PRT;    1140 AA.
AC   054587;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase (EC 6.4.1.1)
GN   PYC OR CGL0689.
OS   Corynebacterium glutamicum (Brevibacterium flavum).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=1718;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=21253;
RA   Koffas M.A.G., Ramamoorthi R., Pine W.A., Sinskey A.J.,
RA   Stephanopoulos G.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 13032;
RA   Peters-Wendisch P.G., Kreutzer C., Kalinowski J., Patek M., Sahm H.,
RA   Eikmanns B.J.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA   Nakagawa S.;
RT   "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- COFACTOR: BIOTIN (BY SIMILARITY).
DR   EMBL; AF038548; AAB92588.1; -.
DR   EMBL; Y09548; CAA70739.1; -.
DR   EMBL; AP005276; BAB98082.1; -.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Biotin; Ligase; Pyruvate; Complete proteome.
SQ   SEQUENCE 1140 AA; 123102 MW; 2A6D4B4ED2FEB531 CRC64;

Query Match          98.2%;    Score 5759;   DB 16;   Length 1140;
Best Local Similarity 99.4%;   Pred. No. 1.7e-299;
Matches 1133;   Conservative    5;   Mismatches   2;   Indels    0;   Gaps    0;

Qy    18 VSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 77
         ::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60
```

RESULT 3

Q8FRQ0
ID   Q8FRQ0   PRELIMINARY;   PRT;   1168 AA.
AC   Q8FRQ0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase (EC 6.4.1.1).
GN   PYC OR CE0709.
OS   Corynebacterium efficiens.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=152794;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA   Kawarabayasi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA   Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
RA   Usuda Y., Sugimoto S.;
RT   "The entire genomic sequence of Corynebacterium efficiens YS-314.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP005216; BAC17519.1; -.
KW   Ligase; Complete proteome.
SQ   SEQUENCE 1168 AA; 126245 MW; A5D5A4DD4DF285F8 CRC64;

Query Match   90.78%; Score 5322; DB 16; Length 1168;
Best Local Similarity 90.8%; Pred. No. 4.1e-276;
Matches 1041; Conservative 46; Mismatches 57; Indels 2; Gaps 2;

Qy   12 KGIITLVSTHTSSTLPAFKKILVANRGEIAVRAFRALETGAATVAIYPREDRGSFHRSF   71
          || ||| || | |||||||||||||||||||||||||| :|:|:|||||||||||||||
Db   25 KGTKTVVTT-IPSTLPAFKKILVANRGEIAVRAFRAAYETGAATVAIYPREDRGSFHRSF   83

Qy   72 ASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFI 131
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   84 ASEAVRIGTEGSPVKAYLDIDEIINAAKKVKADAVYPGYGFLSENAQLARECAENGITFI 143

Qy   132 GPTPEVLDTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGG 191
          ||||||||||||||| ||||||||||||||||| |||||||||||||||||||||||||
Db   144 GPTPEVLDTGDKSKAVSAAKKAGLPVLAESTPSTDIDEIVKSAEGQTYPIFVKAVAGGG 203

Qy   192 GRGMRFVSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVHL 251
          |||||| |:|||||| || | |||:|| ||||||||||||||||||| || :|:|| ||
Db   204 GRGMRFVEKPEDLRELAREASREAEAAFGDGSVYVERAVIKPQHIEVQILGDHTGDVIHL 263

Qy   252 YERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNH 311
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||:|| ||
Db   264 YERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCKSIGYQGAGTVEFLVDEAGNH 323

Qy   312 VFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATIKELGLHTQDKIKTHGAALQCRITT 371
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   324 VFIEMNPRIQVEHTVTEEVTSVDLVKAQMHLAAGATLKELGLLTQDKITTHGAALQCRITT 383

Qy   372 EDPNNGFRPDTGITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVA 431
          ||| :||||| || |||||||||||||||||| |||||||||||||||||||||||||:
Db   384 EDPSNNFRPDTGVITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVS 443

Qy   432 RAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQQAPPADDEQGRIL 491
          ||||||||| |||||||||||||||||||||| :|| |||||||| ||||||||||||||
Db   444 RAQRALAEFNVSGVATNIGFLRALLREEDFTKRRIDTGFIGSHQHLLQQAPPADDEQGRIL 503

Qy   492 DYLADVTVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDAL 551
          :||||||||||||||||| || :||| |||||||||||||||||| ||| || ||||||||
Db   504 EYLADVTVNKPHGERP-ETARPIEKLPEVENIPLPRGSRDRLKQLGPEGFARDLREQDAL 562

Qy   552 AVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDP 611
          ||||||||||||||||||||||||| |||| ||||||||||||||||||||||||||||
Db   563 AVTDTTFRDAHQSLLATRVRSFALTPAARAVAKLTPELLSVEAWGGATYDVAMRFLFEDP 622

Qy   612 WDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDV 671
          | |||||||||||||||||||||||||||||||||||||||||| ||||||||||||:
Db   623 WARLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVQEAKSGVDIFRIFDALNDI 682

Qy   78 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 137
          |||||||||||||||||||| |||||||| ||||||||||||||||||||||||||||||
Db   61 IGTEGSPVKAYLDIDEIINAAKKVKADAVYPGYGFLSENAQLARECAENGITFIGPTPEV 120

Qy   138 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRF 197
          |||||||| ||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   121 LDLTGDKSKAVSAAKKAGLPVLAESTPSTDIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180

Qy   198 VSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVHLYERDCS 257
          |:|||||||||||||||:|||||||||||||||||||||||||||||:||:||||||||||
Db   181 VASPDELRKLATEASREAEAAFGDGAVVYVERAVINPQHIEVQILGDHTGEVHLYERDCS 240

Qy   258 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 317
          |||||||||||||||||||||||||||||||| |||||||||||||||||| |||||||||
Db   241 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCKSIGYQGAGTVEFLVDEKGNHVFIEMN 300

Qy   318 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLHTQDKIKTHGAALQCRITTEDPNNG 377
          ||||||||||||||:||||||||||||||||||||| |||||| |||||||||||||||||
Db   301 PRIQVEHTVTEEVTSVDLVKAQMHLAAGATLKELGLLTQDKIKTHGAALQCRITTEDPNNG 360

Qy   378 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 437
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db   361 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVSRAQRAL 420

Qy   438 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQQAPPADDEQGRILDYLADV 497
          |||.|||||||||||||||||||||| :||:|||| | ||||||||||||||||:||||||
Db   421 AEFNVSGVATNIGFLRALLREEDFTKRRIDTGFIADHPHLLQQAPPADDEQGRILEYLADV 480

Qy   498 TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 557
          ||||||||||||| || :||| |||||||||||||||||| ||| || |||||||||||||
Db   481 TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPEGFARDLREQDALAVTDTT 540

Qy   558 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 617
          |||||||||||||||||| ||||| ||||||||||||||||||||||||||||||||| ||
Db   541 FRDAHQSLLATRVRSFALTPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWARLDE 600

Qy   618 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 677
          |||||||||||||||||||||||||||||||||||| ||:|||||||||||||| |||
Db   601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 660

Qy   678 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLYLKMAEEIVKSGAHILAIKDMAGLLR 737
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLYLKMAEEIVKSGAHILAIKDMAGLLR 720

Qy   738 PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 797
          |||||||||||||||||||||||||||||:||| |||||||||||||||||||||||||||
Db   721 PAAVTKLVIALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 780

Qy   798 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 857
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   781 LSAIVAAFAHTRRDFGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840

Qy   858 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 917
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900

Qy   918 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA 977
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA 960

Qy   978 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREEFYGLIVEGRETLIRLPDVRTP 1037
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREEFYGLIVEGRETLIRLPDVRTP 1020

Qy   1038 LLVRLDAISEPDDKGMRNVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1097
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 LLVRLDAISEPDDKGMRNVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080

Qy   1098 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVVPAATKVEGGDLIVVVS 1157
          ||||||||||||||||||||||||||||||||||||||||||| | |||||||||||||
Db   1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS 1140

```
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Biotin.
SQ   SEQUENCE   1139 AA;  123068 MW;  BA702313451 9FAAA CRC64;

Query Match             90.3%;
Best Local Similarity   90.9%;
Matches 1035;           Conservative

   Score 5297.5;  DB 2;  Length 1139;
   Pred. No. 8.1e-275;
44;  Mismatches 58;  Indels

   Indels    1;  Gaps     1;

Qy       20  THTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIG    79
                ||||||||||||||||||||||||| ||||||||||| :|||||||||||||||||||||
Db        3  TTTPSTLPAFKKILVANRGEIAVRAFRAAYETGAATVAIYPREDRGSFHRSFASEAVRIG    62

Qy       80  TEGSPVKAYLDIDEIIGAAKKVKADAIYPGVFLSENAQLARECAENGITFIGPTPEVLD   139
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       63  TEGSPVKAYLDIDEIINAAKKVKADAVYPGVFLSENAQLARECAENGITFIGPTPEVLD   122

Qy      140  LTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVS   199
             |||||||||||||||||||||||||| :||||||||||||||||||||||||||||||||
Db      123  LTGDKSKAVSAAKKAGLPVLAESTPTDIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVE   182

Qy      200  SPDELRKLATEASREAAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQ   259
             |::|: ||| |||||:| |||||||||||| |||||||||||:||||||||||||||||
Db      183  KPEDLRELARRASREAAAFGDGSVVVERAVIKPQHIEVQLGDHTGDVIHLYERDCSLQ   242

Qy      260  RRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPR   319
             |||||||||||||||||||||||||||||||||||||||||:|:|||||||||||||||
Db      243  RRHQKVVEIAPAQHLDPELRDRICADAVKFCKSIGYQGAGTVEFLVDEAGNHVFIEMNPR   302

Qy      320  IQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFR   379
             |||||| ||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db      303  IQVEHTVTEEVTSVDLVKAQMHLAAGATLKELGLTQDKITTHGAALQCRITTEDPSNNFR   362

Qy      380  PDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAE   439
             |||| |||||||||||||||||||||||||||||||||||||||:||||||:||||||||
Db      363  PDTGVITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVSRAQRALAE   422

Qy      440  FTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTV   499
             |:| ||||||||||||||||||||||||: |||||| |||||||||||:|||||||||||
Db      423  FNVSGVATNIGFLRALLREEDFTKKRIDTGFIGSHQHLLQAPPADDEQGRILEYLADVTV   482

Qy      500  NKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTTFR   559
             |||| || ||:: :||||:|:|:||||||||||||||| |||||||||||||||||||||
Db      483  NKPHGERP-ETARPIEKLPEVENIPLPRGSRDRLKQLGPEGFARDLREQDALAVTDTTFR   541

Qy      560  DAHQSLLATRVRSFALKPAAERAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELR   619
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      542  DAHQSLLATRVRSFALTPAARAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWARLDELR   601

Qy      620  EAMPNVNIQMLLRGRNTVGYIPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAID   679
             ||||||||||||||||||||||||||||||||| || ||||||||||||||||||||||
Db      602  EAMPNVNIQMLLRGRNTVGYIPYPDSVCRAFVQEAAKSGVDIFRIFDALNDISQMRPAID   661

Qy      680  AVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPA   739
             ||||| ||||||||||||||| ::||||||||||||||| ||:|||||||||||||||:|
Db      662  AVLETGSVAEVAMAYSGDLSNPGEKLYTLDYYLNLAEQIVDSGAHILAIKDMAGLLRRA   721

Qy      740  AVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLS   799
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db      722  AAPKLVTALRREFDLPVHVHTHDTAGGQLATYLAAANAGADAVDAASAPLSGTTSQPSMS   781

Qy      800  AIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSN   859
             |:||||||||||||||| ||:||||||||||||||||||||||||||||||||||||||
Db      782  ALVAAFAHTRRDTGLNLQAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSN   841

Qy      860  LRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAA   919
             ||||| ||||||||||||||:|||||||||||||||||||||||||||||||||| ||||
Db      842  LRAQAVALGLADRFELIEDYYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVSPEDFAA   901

Qy      920  DPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSECKAPLTEVPEEEQAHLDADD   979
             |||||||||||||||||||||||||||||||||||||||| |||||| |||||||||:||
Db      902  DPQKYDIPDSVIAFLRGELGTPPGGWPEPLRTRALEGRSQGKAPLAEIPAEEQAHLDSDD   961
```

```
Qy      672  SQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLKMAEEIVKSGAHILAIKD   731
             |||||||||||| ||||||||||||||||| ||||||||||| :||:|| ||||||||||
Db      683  SQMRPAIDAVLETGTSVAEVAMAYSGDLSNPGEKLYTLDYLNLAEQIVDSGAHILAIKD   742

Qy      732  MAGLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSG   791
             |||||| ||:|||||||||||||||||||||||||||||| |||||||||||| |||||
Db      743  MAGLLRPAAATKLVTALRREFDLPVHVHTHDTAGGQLATYYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAG   802

Qy      792  TTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHE   851
             ||||| || |||||||||||||||| ||:||||||||||||||||||||||||||||||
Db      803  TTSQPSMSALVAAFAHTRRDTGLNLQAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHE   862

Qy      852  IPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAG   911
             |||||||||||| ||||||||||||||||:|||||||||||||||||||||||||||||||
Db      863  IPGGQLSNLRAQAVALGLADRFELIEDYYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAG   922

Qy      912  VDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEE   971
             |:| || ||||||||||||||||||||||||||||||||||||||||:||||| ||| |
Db      923  VSPEDFAADPQKYDIPDSVIAFLRGELGTPPGGWPEPLRTRALEGRSQGKAPLAEIPAEE   982

Qy      972  QAHLDADDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRL  1031
             ||||| ||| |  ||| ||:|||||||||||||||||:|||||||||:|||||||||||||
Db      983  QAHLDSDDSAERRGTLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLKEGREELIRL  1042

Qy     1032  PDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRYRDRSVESVTATAEKADSSNKGH  1091
             || ||:| |:|||||:||||||||:|||||||||:|:|||||:||||::||||||:||||
Db     1043  TGVSTPMVVRLDAVSEPDDKGMRNVVVNVNGQIRPIKVDRSVESVTAEKADATNKGH  1102

Qy     1092  VAAPFAGVVTVTVAEGDEVKAGDAVAIIEAMKMEATITPASVDGKIERVVVPAATKVEGGD  1151
             |||||||||||||||||: |||| |||||||||||||:||||||||||||||||||||||
Db     1103  VAAPFAGVVTVTVAEGDEIKAGDAVAIIEAMKMEATITAPVDGVIDRVVVPAATKVEGGD  1162

Qy     1152  LIVVVS  1157
             ||||||
Db     1163  LIVVVS  1168
```

```
RESULT 4

Q8RQL2
ID   Q8RQL2             PRELIMINARY;     PRT;  1139 AA.
AC   Q8RQL2;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase.
GN   PYC.
OS   Corynebacterium efficiens.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=152794;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Akiyoshi N., Nonaka G., Kimura E., Kawahara Y., Sugimoto S.;
RT   "Corynebacterium efficiens pyruvate carboxylase (pyc) gene, complete
RT   CDS.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- COFACTOR: BIOTIN (BY SIMILARITY).
DR   EMBL; AB083299; BAB88903.1; -.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPSase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
```

```
Qy   980 SKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLFPDVRTPLL 1039
         | ||| |||||||||||||||||||||||||||||||||| |||| ||| | ||:::
Db   962 SAERRGTLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLKEGREELIRLTGVSTPMV 1021

Qy  1040 VRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVIATAEKADSSNKGHVAAPFAGV 1099
         ||||| :|||||||||||| || ||||||||||||||||::||||||||:||||||||||
Db  1022 VRLDAVSEPDDKGMRNVVVNVNGQIRPIKYRDRSVESVTATAEKADATNKGHVAAPFAGV 1081

Qy  1100 VTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVVPAATKVEGGDLIVVVS 1157
         |||||||||||||||||||||||||||||||||| ||| |||||||||||||||||||
Db  1082 VTVTVAEGDEIKAGDAVAIIEAMKMEATITAPVDGVIDRVVVPAATKVEGGDLIVVVS 1139
```

RESULT 5
Q9RK64
```
ID   Q9RK64        PRELIMINARY;    PRT;    1124 AA.
AC   Q9RK64;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase.
GN   SCO0546 OR SCF11.26c.
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Seeger K.J., Harris D.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=97000351; PubMed=8843436;
RA   Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA   Kinashi H., Hopwood D.A.;
RT   "A set of ordered cosmids and a detailed genetic and physical map for
RT   the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL   Mol. Microbiol. 21:77-96(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=21996410; PubMed=12000953;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
DR   EMBL; AL939105; CAB59603.1; -.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Complete proteome.
SQ   SEQUENCE 1124 AA;   121071 MW;   1FEF4C4643954D31 CRC64;
```

Query Match   64.7%;   Score 3797;   DB 16;   Length 1124;
Best Local Similarity   65.2%;   Pred. No. 1.5e-194;
  Matches 735; Conservative 156; Mismatches 227; Indels 10; Gaps 6;

```
Qy    29 FKKILVANRGEIAVRAFRAALETGAAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 88
         |:|||||||||||||:||| | ||::::|||| |  || | |||| || :: ||::|||
Db     2 FRKVLVANRGEIAIRAFRAGYELGARTVAVFPHEDRNSLHRLKADEAYEIGEQGHPVRAY 61

Qy    89 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDITGDKSRAV 148
         | :| :| : | ||||:|||| ||:|||::||  | | :|||| :|::||:||||||||
Db    62 LSVEEIVRAARRAGADAVYPGYGFLSENPELARCACEEAGITFVGPSARILELTGNKARAV 121

Qy   149 TAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDELRKIA 208
         |||||  ||||||||||||||:||||:||||||||||||||||||| ::|:
Db   122 AAAREAGVPVLGSSAPSTDVDELVRAADDVGFPVFVKAVAGGGRGMRVERPAQLREAI 181

Qy   209 TEASREAEAAFGDGSVYVERAVINPQHIEVQIIGDRTGEVVHLYERDCSLQRRHQKVVEI 268
         ||| |:|||:|| | :|||| ||||:||| | ||| |||:|||:||:|||::|:|||:|
Db   182 EAASREAAASAFGDSTVFLEKAVVEPRHIEVQILADGEGDVIHLFERDCSVQRRHQKVIEL 241

Qy   269 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHHVTE 328
         ::|| ||: ||:|| |||||| :|| :|::|||||||||:|||||||||||||||||||
Db   242 APAPNLDPALRERICADAVNFARQIGYRNAGTVEFLVDRDGNHVFIEMNPRIQVEHHVTE 301

Qy   329 EVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTTTAY 388
         ||||||||:|:|||:|||||:|||||:|||| | | ||:|||||| ||||||:|||:|||
Db   302 EVTDVDLVQSQLRIAAGQTLADLGLAQENITLRGAALQCRITTEDPANGFRPDTGQISAY 361

Qy   389 RSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 447
         |||||:|:|| |  |||||:||||||||||| |:|:||| ||||||||||||||||||||
Db   362 RSPGGSGIRLDGGTTHAGTEISAHFDSMLVKLSCRGRDFTTAVNRARRAVAEFRIRGVAT 421

Qy   448 NIGFLRALLREEDFTSKRIATGFIGDRPHLLQAPPADDEQGRILDYLADVTVNKPHGVRP 507
         ||| ::| | | |:|| :| ||||| ||||| |  |||||:||||:||||||||||||||
Db   422 NIPFLQAVLDDPDFQAGRVTTSFIEQRPHLLTARHSADRGTKLLTYLADVTVNKPHGERP 481

Qy   508 KDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDAHQSLLA 567
         :: ::| :|||||:||| || | ||:|||||::||||||:|::|||||||||||||||||
Db   482 -ELVDPLTKLPTASAGEPPAGSRQLLAELGPEGFARRLRESSTGVTDTTFRDAHQSLLA 540

Qy   568 TRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVNI 627
         |||:  :| || ||| |:|||||||||| ||||||||||||||:||||||| :||:|:::
Db   541 TRVRTKDMLAVAPVVARTLPQLLSLECWGGATYDVALRFLAEDPWERLAALREAVPNLCL 600

Qy   628 QMLLRGRNTVGYTPYPYDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETNTA 687
         ||||||||||||:|||||  |||||||||:|||||||||||||||:|:|||| ||| |||
Db   601 QMLLRGRNTVGYTPYPYPTEVTDAFVQEAAATGIDIFRIFDALNDVEQMRPAIEAVRQTGSA 660

Qy   688 VAEVAMAYSGDLSDPNEKLYTLDYLYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLVTA 747
         ||||:| |:| :: || |||:|||||||| || || | :|:| ||:||||||| |||:|::
Db   661 VAEVALCYTADLSDPSERLYTLDYYLRLAEQIVNAGAHVLAVKDMAGLLRAPAAATLVSA 720

Qy   748 LRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFAH 807
         ||||:|||||||||||||||||::|| :||||||||  |||| ::|:|||||||||||||
Db   721 LRREFDLPVHLITHDTTGGQLATYLAAIQAGADAVDGAVASMAGTTSCPSLSAIVAATDH 780

Qy   808 TRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSNLRAQATAL 867
         ||||||| | :||||||||||||||||||||||||||||||||||||||||||||||||| 
Db   781 TERPTGLDLQAVGDLEPYWESVRKVYAPFEAGLASPTGRVYHHEIPGGQLSNLRTQAVAL 840

Qy   868 GLADRFELIEDNYAAVNEMLGRPTKVIPSKVVGLALHLVGAGVPDAFAADPQKXDIP 927
         |||||| |||| :|:|||||  | ||||||||||||||||||||||||||:::|||
Db   841 GLGDRFEDIEAMYAAADRMLGRLVKVTPSSKVVGDLALHLVGAGVSPADFEQDPDRFDIP 900

Qy   928 DSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDADDSKERRNSL 987
```

```
Qy   151  AKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGRGMRFVSSPDELRKLATE  210
          :|: |||||| :|||||    || ||    ||||||:|||||||| ::    |
Db   124  ARAAGLPVLSSSEPSSSVDEIMAAADMEFPLFVKAVSGGGRGMRRVTDRESLAEAIEA   183

Qy   211  ASREAEAAFGDGSVYVERAVINPQHIEVQIIGDRTGEVHLYERDCSLQRRHQKVVEIAP  270
          ||||||:|||:||| :: |:|||:|||||::| ::| |:|:|||||:||||||||||:||
Db   184  ASREAESAFGDASVYLEQAVLNPRHIEVQLLADGAGNVMHLFERDCSVQRRHQKVVELAP  243

Qy   271  AQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEEV  330
          |::| ||||::|||||| | | ||| ||||||:|:||:|:|||||:|||||||||||||:
Db   244  APNLSDELRQQICADAVAFARQIGYSCAGTVEFLLDERGHHVFIECNPRIQVEHTVTEEI  303

Qy   331  TEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAYRS  390
          |:|||| :|:|:||||:|  ||| :|::   |||:||||||| | ||||||||:|||||
Db   304  TDVDLVSSQLRIAAGETLADLGLSQDRLVVRGAAMQCRITTEVPANGFRPDTGRITAYRS  363

Qy   391  PGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIG  450
          |||||:||||:: || ||:|||||||||:|||| ||:|||:|||| |||||  ||::|:
Db   364  PGGAGIRLDGGTNLGAEISAHFDSMLVKLTCRGRDFSAAASRARRALAEFRIRGVSTNIP  423

Qy   451  FLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTNKPHGVRPKDV   510
          || |::  | |::| | |:|| |:|||||:: | | | | ||:|||||:|||||| |:|
Db   424  FLQAVIDPDFRAGRVTTSFIDDRPHLLTSRSPADRGTRILNYLADITVNKPHGERPSTV   483

Qy   511  AAPIDKLPNIKDL--PLPRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDAHQSLLAT  568
          : ||||| :::|    :  ||::||::|||::|| | || :|   |:||||||||||||
Db   484  -YPQDKLPPL-DLQAPPPAGSKQRLVELGPEGFAGWLRESKAVGVIDTTFRDAHQSLLAT  541

Qy   569  RVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVNIQ  628
          |||: ||:||:: : :|||||||:| |||||||||||||:|||:|||  || ||||:||
Db   542  RVRTTGLLMVAPYVARSMPQLLSIECWGGATYDVALRFLKEDPWERLAALRESVPNICLQ  601

Qy   629  MLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVLENTAV   688
          |||||||||||||||:|| ||||:||| :|:|||||||||||:|::|||||| | |:|
Db   602  MLLRGRNTVGYTPYPELVTSAFVEEAAATGIDIFRIFDALNNVESMRPAIDAVRETGSTI  661

Qy   689  AEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLVTAL  748
          |||||:|:|||||| |:|||||||||:||:||::|||:|||||||||||||:| ||| ||
Db   662  AEVAMCYTGDLSDPAENLYTLDYYLKLAEQIVEAGAHVLAIKDMAGLLRPAAHTLVSAL  721

Qy   749  RREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFAHT  808
          | ||||||||||||| |||||| |||::||||||||||||::|||||| ||:|||| |||
Db   722  RSRFDLPVHVHTHDTPGGQLATYLAAWSAGADAVDGASAPMAGTTSQPALSSIVAAAHT   781

Qy   809  RRDTGLSLEAVSDLPYWEAVRGLLYLPFESGTPGPTGRVYRHEIPGGQLSNLRAQATALG  868
          : |||| | || || ||||| :: :||||| ||||||||| |||||||||||| ||:|||
Db   782  QYDTGLDLRAVCDLEPYWEAVRKVYAPFESGLPGPTGRVYTHEIPGGQLSNLRQQAIALG  841

Qy   869  LADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFRADPQKYDIPD  928
          | |||| ||:||||  :||||:|||||||||||||||:||||| :  :| ||:|||||||
Db   842  LGDRFEIEANYAADRVLGRLVKVTPSSKVVGDLALALVGAGITAEEFAEDPAKYDIPD   901

Qy   929  SVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDADDSKERRNSLN   988
          |||:||||||| |||||||||||:||:||  :      |: : | | ||| :::::|||
Db   902  SVIGFLRGELGDPPGGWPEPLRTKALQGRGPAR-PVEKLTADDEALL-AQPGPKRQAALN  959

Qy   989  RLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISEP  1048
          ||||| ||| | || |:| :||:|  ::|||||  |  :  :| ||  :: ||:| :|||
Db   960  RLLFPGPTAEFEAHRETYGDTSSLSANQFFYGLRYGEEHRVQL-ERGVELLIGLEAISEA  1018

Qy  1049  DDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAGVTVTVVAEGD  1108
          |::||| |: ||||::||:|||||:: | | ||:||::    | ||||||| : | ||||
Db  1019  DERGMRTVMCIINGQLRPVLVRDRSIASEVPAAEKADRNNADHIAAPFAGVVPAATKVEGGD 1078

Qy  1109  EVKAGDAVAIIEAMKMEATITASVDGKIERVVPAATKVEGGDLIVVVS  1157
          .:  ||:|| ||||||||:||:  : ::  || | | ||||||:|||
Db  1079  SVDAQTIATIEAMKMEAAITAPKAGTVARVAVAATAQVEGGDLLVVVS  1127
```

```
Db   901  DSVVGFLRGELGTPPGGWPEPFRSKALRGRAEAR-PLAELSEDDRDGL---GKDRRATL  955
          ::|||||||||| |||||||| ||   |||: |  ||| |:::|| :|    || |  |
Qy   988  NRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRL-PDVRTPLLVRLDAIS  1046
          |||||| |:||| ||:: || |:| | |:|||||  ||| ||  | |:|  |::|: |:|
Db   956  NRLLFPGPAREFDTHRASYGDTSILDSKDFFYGLRPGKEYTVDLDPGVR--LLIELQAVG  1013
Qy  1047  EPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAGVTVTVVAE  1106
          : |||:||| |::|:||::|:: |||:| :   ||:||: |  |||||||||| | ||||
Db  1014  DADDRGMRTVMSSLNGQLRPIQVRDRSAATDVPVTEKADRANPGHVAAPFAGVVTLAVAE  1073
Qy  1107  GDEVKAGDAVAIIEAMKMEATITASVDGKIERVVPAATKVEGGDLIV  1154
          |||| ||: |:| |||||||:||:  ::  :| :  | |:|||||:|
Db  1074  GDEVEAGATVATIEAMKMEASITAPKSGTVTRLAINRIQQVEGGDLLV  1121
```

RESULT 6
Q9F843
ID   Q9F843   PRELIMINARY;   PRT; 1127 AA.
AC   Q9F843;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase (EC 6.4.1.1).
GN   PYC.
OS   Mycobacterium smegmatis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1772;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20374587; PubMed=10913817;
RA   Mukhopadhyay B., Purwantini E.;
RT   "Pyruvate carboxylase from Mycobacterium smegmatis: stabilization,
RT   rapid purification, molecular and biochemical characterization and
RT   regulation of the cellular level.";
RL   Biochim. Biophys. Acta 1475:191-206(2000).
CC   -!- COFACTOR: BIOTIN (BY SIMILARITY).
DR   EMBL; AF262949; AAG30411.1; -.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   InterPro; IPR000634; S/T_dehydrtse.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMS; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
DR   PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW   Biotin; Ligase.
SQ   SEQUENCE 1127 AA; 120765 MW; 18C132C48425C67B CRC64;

Query Match   64.6%;   Score 3786.5;   DB 2; Length 1127;
Best Local Similarity 66.2%; Pred. No. 5.4e-194;
Matches 747; Conservative 142; Mismatches 233; Indels 7; Gaps 6;

```
Qy   31   KILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLD  90
          |:|||||||||:|||||| | |:ATVA: | || ||:||: | ||  ||  | ||:|||
Db    4   KVLVANRGEIAIRAFRAAYEMGIATVAVPYEDRNSLHRLKADESYQIGEVGHPVRAYLS   63

Qy   91   IDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVTA  150
          :|||| .||: ||||:|||||||||| :||.:|||.||||:||: |||.|||:| ||: |
Db   64   VDEIIRVAKHSGADAVYPGYGFLSENPDLAAKCAEAGITFVGPSAEVLQLTGNKARAIAA  123
```

RESULT 7
P95127
ID   P95127

```
AC   P95127;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 120.4 kDa protein (Pyruvate carboxylase).
GN   PCA OR RV2967C OR MTCY349.20 OR MT3045.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37Rv;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton J., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z83018; CAB05410.1; -.
DR   EMBL; AE007125; AAK47371.1; -.
DR   HSSP; P24182; 1BNC.
DR   TIGR; MT3045; -.
DR   TubercuList; Rv2967c; -.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005471; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR008891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv-carbox.
DR   InterPro; IPR006634; S/T_dehydrtse.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; Pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
DR   PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   1127 AA;  120422 MW;  84B0A4CC1A23CD90 CRC64;
```

```
Query Match      63.3%;  Score 3713.5; DB 16;  Length 1127;
Best Local Similarity  64.4%;  Pred. No. 4.3e-190;
Matches  730;  Conservative 153;  Mismatches 239;  Indels 11;  Gaps  6;
```

```
Qy     29  FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY   88
               |:|||||||||||||||||||:|| | ||| : :| || || |  | || || |
Db      2  FSKVLVANRGEIAIRAFRAAYELGVGTVAVYPYEDRNSQHRLKADESYQIGDIGHPVHAY   61

Qy     89  LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLLTGDKSRAV  148
               | :|||  |: :  |||||||||||||||   || || || ||:| |:||:|
```

```
Db     62  LSVDEIVATARRAGADAIYPGYGFLSENPDLAAACAAAGISFVGPSAEVELAGNKSRAI  121
                |:|||  ||| ||||||||||||||||  || |  ||||:|| |: ||| |  | ||
Qy    149  TAARKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKLA  208
               | : ||||:| |  | :|  ::: |:| ||||||||||||||||||||:|| | |||
Db    122  AAREAGLPVLMSSAPSASVDELLSVAAGMPFFLFVKAVAGGGGRGMRRVGDIAALPFAI  181

Qy    209  TEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVHLYERDCSLQRRHQKVVEI  268
               |||||||||| ||: :|| || : ||:|||| ||:| | |: || | | ||||| ||
Db    182  EAASREARSAFGDPTVYLEQAVINPRHIEVQILADNLGDVIHLYERDCSVQRRHQKVIEL  241

Qy    269  APAQHLDPELRDRICADAVKFCRSIGYQOGAGTYEFLVDEKGNHVFIEMNPRIQVEHTVTE  328
               ||| |||| || | | |  ||:|| || ||||:|| ||| | || ||||| ||||||||
Db    242  APAPHLDAELRYKMCVDAVAFARHIGYSCAGTYEFLLDERGEYVFIEMNPRVQVEHTVTE  301

Qy    329  EVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTTAY  388
               |: ||:|  | | | ||  | |||| : :|  ||||||||||||||||  ||||| : :|
Db    302  EITDVDLVASQLRIAAGETLEQLGLRQEDIAPHGAALQCRITTEDPANGFRPDTGRISAL  361

Qy    389  RSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATN  448
               |: |||||||| |||| :||: |:|:||||| :|| ||||||||:||||||:||:|||| |
Db    362  RTAGGAGVRLDGSTNLGAEISPYFDSMLVKLTCRGRDLPTAVSRARRAIAEFRIRGVSTN  421

Qy    449  IGFLRALLREEDFTSKRIATGFIGDHPHLQAPPADDEQGRILDYLADVTVNKPHGVRPK  508
               || |:|| ||:|::| ::| | |||||||| | | | |||||:|||||||||:|| ||
Db    422  IPFLQAVLDDPDFRAGRVTTSFIDERPQLLTARASADRGTKILNFLADVTVNNPYGSRPS  481

Qy    509  DVAAPIDKLPNIKDLPL---PRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDAHQS  564
               |: | :|:|:|| ||||   ||:|:|:|:|| | | |||  :|| |:|||||||||||:
Db    482  TI-YPDDKLP---DLDLRAAPAGSKQRIVKLGPEGFARWLRESAAVGVTDTTFRDAHQS  537

Qy    565  LLATRVRSFALKPAAEAVAKLLPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPN  624
               |||||||  |:|||||||:|:| ||:||||||:|||:| ||: :||| |||| |||||
Db    538  LLATRVRTSGLSRVAPYLARTMPQLLSVECWGGATYDVALRFLKEDPWERLATLRAAMPN  597

Qy    625  VNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVLET  684
               |:|:||||||||:||| ||||  | | | |||||:||||| | |  | |||:|: ||:
Db    598  ICLQMLLRGRNTVGYTPYPEIVTSAFVQEATATGIDIFRIFDALNNIESMRPAIDAVRET  657

Qy    685  NTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKL  744
               ||  ||| | | || |::|:|| |||:|:| | |||||||:| |:|||||||||: : :|
Db    658  GSAIAEVAMCYTGDLTDPGEQLYTLDYYLKLAEQIVDAGAHVLAIKDMAGLLRPPAAQRL  717

Qy    745  VTALRREFDLPVHVTHDTAGGQLATYPAAAQGADAVDGASAPLSGTTSQPSLSAIVAA  804
               |:||||:|||||:|:||||||:||:||||| ||||||| ||| || | ||||:|||||
Db    718  VSALRSRFDLPVHLTHDTPGGQLASYVAAWHAGADAVDGAAAPLAGTTSQPALSSIVAA  777

Qy    805  FAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSNLRAQA  864
               |||:|||||||:  |:|| ||:||| ||||:| || ||||| |||||||||||||||||||
Db    778  AAHTEYDTGLSLAVCALEPYWEALRKVYAPFSGLPGPTGRVYHHEIPGGQLSNLRQQA  837

Qy    865  TAIGLADRPELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFPAADPQKY  924
               |||| ||: |: |: | |: |  || |:|:|||:|||:|| |||| || ||  |:|: :||
Db    838  IALGLGDRFEEIEEAYAGADRVLGRLVKVTPTSKVVGDLALALVGAGVSADEFASDPARF  897

Qy    925  DIPDSVIAFLRGELGNPPGGWPEPLRTALEGRSEGKAPLTEVPEEEQAHLDADDSKERR  984
               ||| || |||||| | | ||||||||||| ||:|| ||||| | | | |||:|:||:||
Db    898  GIPESVLGFLRGELGDPPGGWPEPLRTAALAGRGAAR-PTAQLAADDEIALSSVGAK-RQ  955

Qy    985  NSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDA  1044
               |:| |||| ||:||||||||: ||:: | :|||:||||: ||:||||  |||||||:||
Db    956  ATLNRLLFPSPTKEFNEHREAYGDTSQLSANQFFYGLRQGEEHRVKL-ERGVELLIGLEA  1014

Qy   1045  ISEPDDKGMRNVANVNGQIRPMVRDRSVESVTATAEKADSSNKGHVAAPFAGVTVTV  1104
               |:|||:  ::|||||| || |:|:| |: :: |||||||| ||:|||||||||||||
Db   1015  ISEPDERGMRTVMCILNGQLRPVLVRDRSIASAVPAAEKADRGNPGHIAAPFAGVTVGV  1074

Qy   1105  AEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVPAATKVEGGDLIVVVS  1157
               | :||:|||||| | | |||| | |||:|:| | |:  | | |||||:||||
Db   1075  CVGERVGAGQTIATIEAMKMEAFITAPVAGTVERVAVSDTAQVEGGDLLVVVS  1127
```

```
RESULT 8
Q50450
```

```
Db  361 CDPPAVPVSAWTAAP-----TWRRNQPVLRLHAGQADLSGRDLPTAVSRARRAIAEFRIR  415
Qy  444 GVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKPH  503
Db  416 GVSTNIPFLQAVLDDPDFRAGRVTTSFIDERPQLLTARASADRGTKILNFLADVTVNNPY  475
Qy  504 GVRPKDVAAPIDKLPNIKDLPL----PRGSRDRLKLQLGPAAFARDLREQDALAVTDTTFR  559
Db  476 GSRPSTI-YPDDKLP---DLDLRAAPPAGSKQRLVKLGPEGFARWLRESAAVGVTDTTFR  531
Qy  560 DAHQSLLATRVRSFALKPAEAVAKLTPELLSVEAWGGAIYDVAMRFLFEDPWDRLDELR  619
Db  532 DAHQSLLATRVRTSGLSRVAPYLARTMPQLLSVECWGGATYDVALRFLKEDPWERLATLR  591
Qy  620 EAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAID  679
Db  592 AAMPNICLQMLLRGRNTVGYTPYPEIVTSAFVQEATATGIDIFRIFDALNNIESMRPAID  651
Qy  680 AVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLKMAEEIVKSGAHILAIKDMAGLLRPA  739
Db  652 AVRETGSAIAEVAMCYTGDLITDPGEQLYTLDYLKLAEQIVDAGAHVLAIKDMAGLLRPP  711
Qy  740 AVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLS  799
Db  712 AAQRLVSALRSRFDLPVHLHTHDTPGGQLASYVAAWHAGADAVDGAAAPLAGTTSQPALS  771
Qy  800 AIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSN  859
Db  772 SIVAAAAHTEYDTGLSLSAVCALEPYWEALRKVYAPFESGLPGPTGRVYHEIPGGQLSN  831
Qy  860 LRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAA  919
Db  832 LRQQAIALGLGDRFEEIEEAYAGADRVLGRLVKVTPTSKVVGDLALALVGAGVSADEFAS  891
Qy  920 DPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDADD  979
Db  892 DPAREGIPESVLGFLRKGELGDPPGGWPEPLRTAALAGRGAAR-PTAQLAADDEIALSSVG  950
Qy  980 SKERRNSLNRLLFPKPTEEFLEHRRFGNTSALDDREFFYGLVEGRETLIRLPDVTPLL  1039
Db  951 AK-RQATLNRLLFPSPTKEFNEHREAYGDTSQLSANQFFYGLRQGEEHRVKL-ERGVELL  1008
Qy 1040 VRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAGV  1099
Db 1009 IGLEAISEPDERGMRTVMCILNGQLRPVLVRDRSIASAVPAAEKADRGNPGHIAAPFAGV  1068
Qy 1100 VTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVV  1141
Db 1069 VTVGVCVGERVGAGQTIATIEAMKMEAPIFAPVAGTVERVAV  1110
```

```
ID   Q50450        PRELIMINARY;       PRT;   1124 AA.
AC   Q50450;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyc.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Smith D.R.;
RL   Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Robison K.;
RL   Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- COFACTOR: BIOTIN (BY SIMILARITY).
DR   EMBL; U00024; AAA50948.1; -.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   InterPro; IPR000634; S/T_dehydrtse.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; Pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
DR   PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW   Biotin.
SQ   SEQUENCE 1124 AA;  120672 MW;  07E1692E1203491 CRC64;

Query Match             59.4%;  Score 3484;  DB 2;  Length 1124;
Best Local Similarity   62.2%;  Pred. No. 8.2e-178;
Matches   698;  Conservative 151;  Mismatches 251;  Indels 22;  Gaps   9;

Qy   29 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY   88
        | |:|||||||||:||:::::|:|||  |||  || :||| || |||||| : |  | ||
Db    2 FSKVIVANRGEIAIRAFRAAYELGVGTVAVYPYEDRNSQHRLKADESYQIGDIGHPVHAY   61

Qy   89 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV  148
        |:|||||||::||| |||||||||||| || || ::|:||||:|||||| |:|||||:||
Db   62 LSVDEIVATARRAGADAIYPGYGFLSENPDLAAACAAAGISFVGPSAEVLELAGNKSRAI  121

Qy  149 TAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKLA  208
        |||: |||||    ::|:|||| |||:| :|||| ||||||||||||||||||:  |:||
Db  122 AAAREAGLPVLMSSAPSASVDELLSVANGMPFPLFVKAVAGGGGRGMRRVGDIAALPEAI  181

Qy  209 TEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEI  268
        ||||| ||  :|||||:|||| |||||||| ::|:|:||||:||:|:|| ||||:| :::
Db  182 EAASREAEASFGDPTVYLEQAVINPRHIEVQILADNLGDVIHLYERDCSVQRRHQKVIEL  241

Qy  269 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE  328
        |||:||| |||| || ||| |  :||  || |||||||||| | |||||||:| |||||:
Db  242 APAPHLDAELRYKMCVDAVAFARHIGYSCAGTVEFLLDERGEYVFIEMNPRVQVEHTVTE  301

Qy  329 EVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAY  388
        |:|:||||||:|| ||||| ||| |:| |||:|  |||||||||||:| |||  |||:|
Db  302 EITTVDLVASQLRIAAGETLEQLGLRQEDIAPHGAALQCRITTEDPANGFRP-TRAGSAR  360

Qy  389 RSPGGAGVRLDGAAQLGGEITAHFDSMLVKM-----TCRGSDFETAVARAQRALAEFTVS  443
        | |   |||  | ||| || :|:|||||:| :   : ::||:|  |||||||:| |||||
```

```
RESULT 9
Q8JHF6
ID   Q8JHF6        PRELIMINARY;       PRT;   1178 AA.
AC   Q8JHF6;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase (EC 6.4.1.1).
GN   PYC.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22145742; PubMed=12150961;
RA   Jitrapakdee S., Nezic M.G., Cassady A.I., Khew-Goodall Y.,
RA   Wallace J.C.;
RT   "Molecular cloning and domain structure of chicken pyruvate
RT   carboxylase.";
RL   Biochem. Biophys. Res. Commun. 295:387-393 (2002).
```

```
DR   EMBL; AF509529; AAM92771.1; -.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; Biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
KW   Ligase.
SQ   SEQUENCE   1178 AA;   127253 MW;   F0722FEAB8BF39A5 CRC64;

Query Match          43.5%;
Best Local Similarity   47.5%;   Pred. No. 7.6e-128;
Matches  559;  Conservative  168;  Mismatches  416;  Indels  33;  Gaps  15;

Qy     7 GGLLLKGIITLV------STHDISSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIY  59
           :||: |||                ::||: | ||||||||||||:||:  |: :::
Db     9 GGRALLGAWRLPLLRPPPGSVRSASCQP-IRKVLVANRGEIAIRVFRACTELGLRTVAVY  67

Qy    60 PREDRGSFPHRSAFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIPYGPGEFLSENAQL 119
           ::||  | | |     :|   ||||| |  ::|| || :||||||:|||||||||
Db    68 SEQDTGQMHRQKADEAYIVGRGLPPVQAYLHVPDIIRVARENAVDAIHPGYGFLSERADF 127

Qy   120 ARECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVL-AESTPSKNIDDIVKSAEGQ 178
           | ||  ||  |:|| |||| |  ||| ||::|||| ||:   :|  |  :|||:||| |
Db   128 AQACVDAGVRFVGCPPPEVVRKMGDKVEARSIA1AAGVPVVPGTSAPVATLGEAQDFAARV 187

Qy   179 TYPIFVKAVAGGGGRGMRFVSSPDELRKLATEASREAEAAFGDGSVYERAVINPQHIEV 238
           : ||:|||| |||||||| | ||:||: : |:| ||||:|||||::::|:::||:|||
Db   188 GFPIIFKAAHGGGGRGMRAVRGPQELEESFRASSEALAAFGDGALFVEKLMERPRHIEG 247

Qy   239 QIIGDRTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGA 298
           |||||: |:|||||||||| |||||||||||||: ||::|| :| :|||:   ::|| |
Db   248 QIIGDKHGNVVHLYERDCSIQRRHQKVVEIAPAARLDFQLRAQLASDAVRIAQQVGYENA 307

Qy   299 GTVEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKI 358
           |||||||| :|:| ||| | |||||||||||:| |||| ||: | |||:| ||||| | :
Db   308 GTVEFLVDRDGKHYFIEVNSRLQVEHTVTEEITGVDLVQAQLLVAAGRSLSELGLQQDSV 367

Qy   359 KTHGAALQCRITTEDPNNGFRPDTGTTAYRSPGGAGVRLDGAAQL-GGEITAHFDSMLV 417
           : ::||:||| |||||: ||:|||| ::: ||| | ||||||::: | |:| ||:||:|
Db   368 RVNGCAIQCRVTTEDPARGFQPDTGRIEVFRSGEGMGIRLDGASAFQGALISPHYDSLLV 427

Qy   418 KMTCRGSDFETAVARAQALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHL 477
           |:    | :| : |:|:|| ||:: ||:|||.||:::| |: |  | :  :|: |: | |
Db   428 KVIAHGPDQPSAAAKMSRALGEFRIRGVKTNIPFLQNVLAHPQFLGGAVDTQFIDENPEL 487

Qy   478 LQAPPADDEQGRILDYLADVTVNKPHGVRP-KDVAAPIDKLPNIKDLPLPRGS-----RD 531
           :  :|   ::| :|  |  | |: |    |  :| |   ::|::: |:| ::|       :
Db   488 FHLRPSQNRAQKLLHYLGHVMVNGPSTPLPVKAKAAVVEPVPP---PVPMGSSPPEGLRA 543

Qy   532 RLKQLGPAAFARDLREQDALAVDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLS 591
            | :|||| ||  ||   || |||||||||||||||||:    : : ::|  ||| |  |
Db   544 VLQREGPAGFARALRGHRGLLLXDTTFRDAHQSLLATRVRTRDLARIAPFVAHSLSPLCS 603

Qy   592 VEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGVTPYPDSVCRAFV 651
           :| |||||:|||||||  :|| || ||| :||:| :|||||| | |||| |||:| : |
Db   604 METWGGATFDVAMRFLHECPWERLRELRRLVPNIPFQMLLRGANAVGVTNYPDNVIYRFC 663

Qy   652 KEAASSGVDIFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDY 711
           : ||::| ||||||||||::  ||  |:|| :| |||  || |||| :|||  | | |||
Db   664 EVAAANGMDIFRIFDALNYLPNLLLGVEAVGRAG-AVVEAALSYTGDVADPTRTKYSLDY 722

Qy   712 YLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVHTHDTAGGQLAT 770
           || :|:|:|::|: |||:||||||| ||| :|:| :||: | |:|:|||||||: |  :||
Db   723 YLGLAKELVEAAGTHILCIKDMAGLLTPAAARLLVSSLRDRFDVPIHVHTHDTRAGAAIAT 782

Qy   771 YFAAAOAGADAVDGASAPLSGTTSQPSLSAIVAAFAHTRDTGLSLEAVSDLEPYWEAVR  830
           :|||  ||| | || : |:|||||||| ||||| :| |  || | : :   :||| |:|
Db   783 LLAAANADADVDVAVDAMSGMTSQPSMGAIVACARGTPLDTGIALERVFEYSEYWEGAR  842

Qy   831 GLYLPFESGTPGPTGR--VRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLG  888
           |||  | : ::   | |   :||||||  ||  ||: :|| :  | |:::  ||:|: |
Db   843 GLYAAFDCTATMKSGNADVYENEIPGGQYTNLHFQAHAMGLGHKFKEVKKAYAEANKLLG  902

Qy   889 RPTKVTPSSKVVGDLALHLVGAGVDPADFRADPQKYDIPDSVIAFLRGELGNPPGGWPEP  948
           |  |||||||||||||  :|  |:| | :  |  :  :  ||:||| | :| ||||:|||
Db   903 DLIKVTPSSKVVGDLAQFMVQNGLSREEAEARADELSFPLSVVEFLQGYIGTPPGGFPEP  962

Qy   949 LRTRALEG--RSEGKAPLTEVPEEEQA---HLDADDS--KERRNSLNRLLFPKPTEEFLE 1001
            |::    |  ||   || |:| :| |    :|  | |   | | ||  |: |  :: |
Db   963 FRSKVIKDLPKVEGRPGASLPPLDFEALSQELGARDGTPPSPEDLLSAALYPKVYAEFRD 1022

Qy  1002 HRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVN 1061
           ::::||||| : ||:  :|  || | |: : |:    |: |  | | ::    :  | :
Db  1023 FTSTFGPVSCLGTRLFLEGPTIAEFFEVELERGKT-LHIKALAGDLNAAGQREAFFELN 1081

Qy  1062 GQIRPMRVRDRSVESVTATEAKADSNKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIE 1120
           || |::||||  : : :    |||:| ||| || || ||:|  :| : |:||  :  :
Db  1082 GQLRSILVRDTQALKEMHVHPKADRSAKGQVGAPMPGEVVRVKEGEAVEKGAPLCVLS 1141

Qy  1121 AMKMEATTIRASVDGKIERVVVPAAIKVEGGDLIVVV 1156
           ||||| ::  |  || :  | : | || ||||:|  :
Db  1142 AMKMETVVTAPRGGTVSRLHVRPGMSLEGDLIIAEI 1177


RESULT 10
Q9K9M0
ID   Q9K9M0      PRELIMINARY;      PRT;   1150 AA.
AC   Q9K9M0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase.
GN   PYCA OR BH2625.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
DR   EMBL; AP001516; BAB06344.1; -.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   1150 AA;   128858 MW;   ED6788BE8A9F3BD4 CRC64;

Query Match          43.2%;   Score 2535.5;  DB 16;  Length 1150;
```

```
Best Local Similarity  46.9%;  Pred. No. 4.6e-127;
Matches  540;  Conservative 185;  Mismatches 399;  Indels  27;  Gaps  13;

Qy    26 LPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPV   85
            :|:||||||||||| ||| :|| :: ||| : |  |  ::|| ::|| |||||||
Db     4 LKNIKKVLVANRGEIAIRIFRACELHIRTVAIYSKEDTGAYHRYKADEAYLVGEGKKPI   63

Qy    86 KAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKS  145
         :||||| : ||  |:  |:||| ||||||||| : |:|| :||:||| ||: | ||||:
Db    64 EAYLDIEGIIEIAKRHGVDAIHPGYGFLSENIEFAKRCHEEGIIFIGPELEHLYMFGDKV  123

Qy   146 RAVTAAKKAGLPVLAEST-PSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDEL  204
          | : |:|| |||: : :  |:: :::| :|| : ||::|||| ||||||||||:| ::::
Db   124 QAREQAIKANLPVIPGSDGPVSSLEDVRAFADKHGYPFIIKAALGGGGRGMRIVRSENDV  183

Qy   205 RKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQK  264
          : :  | ::|:|||| :  |||: .|:|:|||||||.|: |: :|||||||||:|||||||
Db   184 QESYERAKSEAKAAFGNDEVYVEKFIENPKHIEVQILADKHGNTLHLYERDCSVQRRHQK  243

Qy   265 VVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEH  324
         |||:||: .| :::|:||| :||:  : ||:  |||||||||:.||: :||:|||||||||
Db   244 VVEVAPSVSLSEDVRERICQAAVQLAENVNYVNAGTVEFLVDREGNFYFIEVNPRIQVEH  303

Qy   325 TVTREVTEVDLVKAQMRLAAGATL--KELGL-TQDKIKTHGAALQCRITTEDPNGFFRPD  381
          | .: || :|:|::  :| |  |   ::|:|  :::|  |||| |  :|||||:.||: |
Db   304 TITEMVTGIDIVQSQLFIADGEHLHGDRLGIPKQEEIVCHGYAIQSRVTTEDPSNGFLPD  363

Qy   382 TGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEF  440
         || | ||||..| |||||  .. | | || ::||:|||:|  ::|| | |||:||||:|:
Db   364 TGRINAYRSGGGFGVRLDAGNGFQGAVITPYYDSLLVKVSTWALTFEGAAAKMLRNLREF  423

Qy   441 TVSGVATNIGFLRALLREDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDLYILADVTVN  500
          : |: |||.|| : :: ::| |  | |:|   |:|: :|| ::. |    ::|||:|:||
Db   424 RIRGIKTNIAFLENVVQHRQFLSGEYNTSFIDQTPELFVFPKKRDRGTKMLSFIGETIVN  483

Qy   501 KPHGVRPKDVAAPIDK--LPNIK-DLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  557
          | |:  :|  .|:||   :|::| |   | | ::: ::||| |:: :|:: : | :|||||
Db   484 GYPGLE-KTKKPVFDKPKPVPKLKLSEPIPDGTKQILDQHGPEGLAKWVKEQKHVLLTDTT  542

Qy   558 FRDAHQSLLATRVRSFALKPAAFAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  617
         |||||||||||||:. |||.| :  .|| || :| |:||||:|||||||| |||| || :
Db   543 FRDAHQSLLATRVRTHDLKQIAEPTARLLPNLFSMEMWGGATFDVAMRFLHEDPWERLLI  602

Qy   618 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  677
         || . ||| :||||| .| ||| .||:.: | ||: : :.|:|:||||| ||:|  |: |
Db   603 LRKKAPNVLFQMLLRASNAVGYKNYPDNLIREFVDKSANAGIDVFRIFDSLNWVEGMKLA  662

Qy   678 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLKMAEEIVKSGAHILAIKDMAGLLR  737
         |:||  || ||| |   |:||: |.. | | |:|.|: ||| | |||||.|||||||||
Db   663 IEAVGEAN-KIAEATICYTGDILDSSRPKYDLAYYKKLAKELEAAGAHILGIKDMAGLLK  721

Qy   738 PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  797
         | || :||| |:: . :|||:|||||:|..: || |:: || |:||.| | :|||||||
Db   722 PEAAYQLVAELKDTVTIPVHLHTHDTSGNGIFTYARAIEAGVDIVDVAVSSMAGLTSQPS  781

Qy   798 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLYPFESGTPGPGTGRVYRHEIPGGQL  857
          .|: |: |. : |:.:.: |:::| | | | | |: :.|||||... ||   | |:|||| |
Db   782 ANSLYYALADSERQPNVNITALEQLAEFWEETRKFYAGFESGMNAPHTEVYEHEMPGGQY  841

Qy   858 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHIVGAGVDPADF  917
         |||  : | ||| |:: |:: .|   |:| |.|||||||||||||:||::| : ::  :
Db   842 SNLQQQAKAVGLGHRWNEVKKMYRTVNDMFGDVVKVTPSSKVVGDMALYMVQNDLTEEEV  901

Qy   918 AADPQKYDIPDSVIAFLRGELGNPPGGWPELRTRALEGRSEGKAPLTEVPEEQAHLDA  977
            .|:||:||| ::|:::|| . | |:|::| ::  .:. |  :|:   | |:
Db   902 YENGHKLDFPDSVVEFFQGLGQPYQGFPKKLQEIILKGRK---PITNRPGENMEPIQF  957

Qy   978 DDSKER-RNSLNR------LLFKPTEEFLEHRRFGNTSALDDREFFYGLVEGRE  1026
          : .|  :::| :       : ::  |.|:|||| | ::|| ::||||| |.||
Db   958 EAIKEELYNKLDRQVTSHDILSYALYPKVFMEFERFRQTFGDVSVLDPTFFYGLRPGEE  1017

Qy  1027 TLIRLPDVRTPLLVRLDAISEPDDKGMRNVANVNGQIRPMRVRDRSVESVTAERKADS  1086
         :|: |||:||||:||| || |  : .||||||:||:| .|| |||::||:|:| |.|

Qy  1018 IEVEIEQGKT-LIVKFISLSKPQDDGNRIVTFELNGQPREVLIKDQSVKTSISRPKADK  1076
          :| |: .||  |:|:|| ||||:|:||||||||:||||||||||||||:||||||||:
Db  1087 SNKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVPPAAT  1145

Db  1077 SNPNHIGASMPGTVKALVEKGDKVKQGDHLMITEAMKMETTVQAPFDGEVVALHVKDGD  1136

Qy  1146 KVEGGDLIVVV  1156
          :|: |||:: |
Db  1137 AIQTGDLLIEV  1147
```

RESULT 11
Q9KWU4

```
ID   Q9KWU4;          PRELIMINARY;        PRT;    1148 AA.
AC   Q9KWU4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase (EC 6.4.1.1).
GN   PYCA.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z99111; CAB13359.1; -.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
```

DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Ligase; Complete proteome.
SQ   SEQUENCE 1148 AA; 127936 MW; 98D14DD56F361620 CRC64;

Query Match   43.1%; Score 2530; DB 16; Length 1148;
Best Local Similarity   46.0%; Pred. No. 9.1e-127;
Matches 533; Conservative 188; Mismatches 390; Indels 48; Gaps 13;

Qy   28 AFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKA 87
               :||||||||||| ||| |||:|| ||:: |:::| || || | |: || |
Db   5 SIQKVLVANRGEIAIRIFRACTELNIRTVAVYSKEDSGSYHRYKADEAYLVGEGKKPIDA 64

Qy   88 YLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRA 147
               :|| : ||| | |||||||||:|| |:|| ||||| |: ||:| |
Db   65 YLDIEGIIDIAKRNKVDAIHPGYGFLSENIHFARRCEEEGIVFIGPKSEHLDMFGDKVKA 124

Qy   148 VTAAKKAGLPVLAEST-PSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRK 206
               |:||:| :||| | ||:||::|| | :||||||||||| ||||||| |:||:|
Db   125 REQAEKAGIPVIPGSDGPAETLEAVEQFGQANGYPIIKASLGGGGRGMRIVRSESEVKE 184

Qy   207 LATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVV 266
               || : | | ||::|| |::: |: :|| ||:|::|| || ||||||| :||||||||
Db   185 AYERAKSEAKAAFGNDEVVVEKLLENPKHIEVQVIGDKQGNVHLFERDCSVQRRHQKVI 244

Qy   267 EIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTV 326
               |:| | :| ||:| || | |||: |:| | |||| |:||||||:|||
Db   245 EVAPSVSLSPELRDQICEAAVALAKNVNYIINAGTVEFLV-ANNEFYFIEVNPRVQVEHTI 303

Qy   327 TEEVTEVDLVKAQMRLAAGATL-KELGLTQDK-IKTHGAALQCRITTEDPNNGFRPDTG 383
               ||||| | ::| | :| :| |: :: | || | |||:|| |:||||:::|| |||
Db   304 TEMITGVDIVQTQILVAQGHSLHSKKVNIPEQKDIFTIGYAIQSRVTIEDPQNDFMPDTG 363

Qy   384 TITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTV 442
               ::|| ||| ||:||| |||:|||||| ||::|:| ||| || ||::| : || |:
Db   364 KIMAYRSGGGFGVRLDTGNSFQGAVITPYYDSLLVKLSTWALITFEQAAAKMVRNLQEFRI 423

Qy   443 SGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQQPPADDEQGRIIDYLADVTVN-- 500
               ||| ||| ||| ||||||| | :||| :| ||| || || ::: |:|| | |:|
Db   424 RGIKTNIPFLENVAKHEKFLTGQYDTSFIDTTPELFNFPKQKDRGTKMLTYIGNVTVNGF 483

Qy   501 ------KPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAV 553
                     ||| | :| ||| ||||:| | ||:| ||| ||||:| :||:||| |
Db   484 PGIGKKEKPAFDKPLGVKVDVDQQP------ARGTKQILDERGAEGLANWVKBQKSVLL 536

Qy   554 TDITFRDAHQSLLATRVRSFALKPAAEAVAKLIPELLSVEAWGGATYDVAMRFLFEDPWD 613
               ||:|||||||||||| ||| | ||||:|| ||| |: ||| |||| ||:||| |:||:
Db   537 TDTTFRDAHQSLLATRIRSHDLKKIANPTAALWPELFSMEMWGGATFDVAYRFLKEDPWK 596

Qy   614 RLDELREAMPNVNIQMLLRGRNTVGYIPYPDSVCRAFVKEAASSGVDFIRFIDALNDVSQ 673
               ||||||:| |:|:||::|||||||||||||| | |: :| :| | |||||| | |||
Db   597 RLEDLRKEVPNTLFQMLLRSSNAVGYTNYPDNVIKEFVKQSAQSGIDVFRIFDSLNWVKG 656

Qy   674 MRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMABEIVKSGAHILAIKDMA 733
               ||| |:|| | || || | ::| || ||||:||:||:|| |||||||| |||||
Db   657 MTLAIDAVRDTG-KVAERAICYTGDILDKNRTKYDLAYYTSMAKELEAAGAHILGIKDMA 715

Qy   734 GLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTT 793
               ||:|||| || | :||: ||| ||||| | ||| | |:|:|: ||| || | |::| |
Db   716 GLLKPQAAYELVSALKETIDIPVHLHTHDTSGNGIYMAKAVEAGVDIDVAVSSMAGLT 775

Qy   794 SQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIP 853
               ||| || |: | | | || || :: ::| :| | | | |||| ||: |||
Db   776 SQPSASGFYHAMEGNDRRPEMNVQGVELLSQYWESVRKYYSEFESGMKSPHTEIYEHEMP 835

Qy   854 GGQLSNLRQATALGLADRFELIEDNYAAVNEMLGRPIKYTPSSKVVGDLALHLVGAGVD 913
               ||| ||| | | || | |: || || |:| | |:| |||||||||:|:|||:| ::
Db   836 GGQYSNLQQAKGVGLGDRWNEVKEMYRRVNDMFGDIVKVTPSSKVVGDMALYMVQNNLT 895

Qy   914 PADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEE-- 970
               |:|: |:| |:: | || : :| || | || |: :: |:| |:||
Db   896 EKDVYEKGESLDFPDSVVELFKGNIGQPHGGFPEKLQKLILKGQE---PITVRPGELLE 951

Qy   971 -------EQAHLDADSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFF 1018
                      :||:|: | ::| : | | | |: | |::| :
Db   952 PVSFEAIKQEFKEQHNLEISD---QDAVAYALYPKVFTDVKTTESYGDISVLDTPTFF 1007

Qy   1019 YGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVVRDRSVESVT 1078
               |||:|| | | ||| :: |:| |: :| : |:|| |:: :||:||| |:| ||
Db   1008 YGMTLGEIEVEIERGKT-LIVKLISGEPQDATRVVYFELNGQPREVVIKDESIKSSV 1066

Qy   1079 ATAEKADSSNKGHVAAPFAGVTVTVAE-GDEVKAGDAVAIIEAMKMEATITASVDGKIE 1137
               ||:| |:| | | | ||:| | |||:||| ::|::||:|| |: :: ||:||
Db   1067 QERLKADRTNPSHIAASMPGTVIKVLAEAGTKVNKGDHLMINEAMKMETTVQAPFSGTIK 1126

Qy   1138 RVVVPAATKVEGGDLIVVV 1156
               :| | | ||:|| |::
Db   1127 QVHVKNGEPIQTGDLLLEI 1145

RESULT 12
Q9DDT1
ID   Q9DDT1   PRELIMINARY;   PRT;   1180 AA.
AC   Q9DDT1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Pyruvate carboxylase.
GN   PC.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21100417; PubMed=11167010;
RA   Yoder J.A., Litman G.W.;
RT   "The zebrafish fth1, slc3a2, men1, pc, fgf3 and cycd1 genes define two
RT   regions of conserved synteny between linkage group 7 and human
RT   chromosome 11q13.";
RL   Gene 261:235-242(2000).
CC   -!- COFACTOR: BIOTIN (BY SIMILARITY).
DR   EMBL; AF295372; AAG37836.1; -.
DR   HSSP; P24182; 1BNC.
DR   ZFIN; ZDB-GENE-000831-1; pc.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPSase_L_D2.
DR   InterPro; IPR005481; CPSase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
KW   Biotin.
SQ   SEQUENCE 1180 AA; 129984 MW; 26118F2691400E24 CRC64;

Query Match   43.0%; Score 2523.5; DB 13; Length 1180;
Best Local Similarity   46.6%; Pred. No. 2.1e-126;
Matches 550; Conservative 179; Mismatches 419; Indels 31; Gaps 16;

Qy   6 LGGLLLKGIITLV-------STHTSS-TLP-AFKKILVANRGEIAVRAFRAALETGAA 54

```
Db   1081 LNGQLRSVLVKDIAAMKEMHFHPKALKDVRGQVGAPMPGKVVEVKVKAGQKVEKGQPLCV 1140

Qy   1119 IEAMKMEATTTASVDGKIERVVVPAATKVEGDLIVVVS 1157
          ||||| :|||     |||   |||||||||||
Db   1141 LSAMKMETVVNSPISGIISKVHVNADSSLEGEDLILEIT 1179
```

Query Match   42.8%; Score 2509; DB 2; Length 1147;
Best Local Similarity   46.8%; Pred. No. 1.2e-125;
Matches 539; Conservative 181; Mismatches 397; Indels 34; Gaps 15;
```

```
Db      5 LGGSRAKTGILLRRAGLFCRSAHLSSQTLEYHP_IKKVWVANRGEIAIRVFRACTELGIR 64
           ||| :  |    ::| |:  | :||      || :     |||  ||
Qy      5 TVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLDIDELIGAAKKVKADAIYPGYGFLS 114
Db     55 TVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLDIDELIGAAKKVKADAIYPGYGFLS 114
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     65 TVAVYSEQDTGQMHRQKADEAYLIGRGLSPVAAYLHIPDIIKVAKENNVDAIHPGYGFLS 124
Db        ||: : | ||   | ||::|     ||::| | | ::|||:| | : ||||:|||||||

Qy    115 ENAQLARECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVL-AESTPSKNIDDIVK 173
Db    125 ERADPAQACAEAGVRFIGPSPEVVRKMGDKVEARALAIKAGVPVVPGTDAPISCLQEAQE 184

Qy    174 SAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKLATEASREAEAAFGDGSVYVERAVINP 233
Db    185 FAKIYDFPIIFKAAYGGGGRGMRVVRNYEELEENYQRAYSEALAAFGNGALFVEKFIEKP 244

Qy    234 QHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSI 293
Db    245 RHIEVQILGDKYGNVIHLYERDCSIQRRHQKVVEIAPAPAHLDSHLRDRLTLDSVNLAKQV 304

Qy    294 GYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGL 353
Db    305 GYENAGTVEFLVDKHGKHYFIEVNSRLQVEHTVTEEITDVDLVHAQLRVCEGRSLPELGL 364

Qy    354 TQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQL-GGEITAHF 412
Db    365 EQDKIQINGCAIQCRVTTEDPSRGFQPDTGRIEVFRSGEGMGIRLDSASAFQGAIISPHY 424

Qy    413 DSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIG 472
Db    425 DSLLVKIASGKDLPTAATKMHRALTEEFRVRGVKTNIPFLQNVLSNSQFLYATVDTQFID 484

Qy    473 DHPHLLQAPPADDEQGRILDYLADVTVNKPHGVRP-KDVAAPIDK-LPNIKDLPLPRGSR 530
Db    485 ENQNLFNLKPTQNRAQKLLHYLGHVMVNGPMTPIPVKAKPSPVDPVIPSVSLGEPPLGFR 544

Qy    531 DRLKQLGPAAFARDLREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEEAVAKLTPELL 590
Db    545 EVLLREGPEGFARAIRQHQGLLLMDTTFRDAHQSLLATRVRTHDLKKIAPYVSHNFSNLF 604

Qy    591 SVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAF 650
Db    605 SLENWGGATFDAAMRFLSECPWKRLQELRALIPNVPFQMLLRGANAVGYTNYPDNAVFKF 664

Qy    651 VKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLD 710
Db    665 CEVAKENGMDIFRVFDSLNYIPNMLLGMEAAGAAG-GVVEAAISYTGDVSDPMRQKYSLD 723

Qy    711 YYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVTHDTAGGQLA 769
Db    724 YYLKLADELVKAGTHILSIKDMAGLLKPQASRLLIEALRDRFPDIPIHVTHDTAGAGVA 783

Qy    770 TYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAV 829
Db    784 AMLACAQAGADIVDVAVDSMAGMTSQPSMGAIVACTKGTKLDTGISLDKVFDYSEYWEVA 843

Qy    830 RGLYLPFESGTPGPTGR--VYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEML 887
Db    844 RGLYAPFDCTATMKSGNADVYENEIPGGQYTNLHFQAHSMGLGNKFKEVKAYTEANKLL 903

Qy    888 GRPTKVTPSSKVVGDLALHLVGAGVDPADFAADPKYDIPDSVIAFLRGELGNPPGGWPE 947
Db    904 GDLIKVTPSSKIVGDLAQFMV-QNLSRAEVEKRADELSPPLSVVEFLQGHGIGHIGFPE 962

Qy    948 PLRTRALEG--RSEGKAPLTEVPEEEQAHLDA-----DDSKERRNSLNRLLFPKPTEEF 999
Db    963 PFRSKVLKSLPRIEGR-PGASLPPMDFEALESGLRAAHGDEITPEDVMSAAMYPKVFQEF 1021

Qy   1000 LEHRRRFGNTSALDDREFFYGLVEGRETILRLPDVRTPLLVRLDAISEPDDKGMRNVVAN 1059
Db   1022 KEFTSTFGPVDCLNTRLFLDGPKIAEEFQVELERGKI-LHIKALALGDLNKSGQREVFFE 1080

Qy   1060 VNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAI 1118
Db        :||||: |::    ||     | |   |: ||  | |   |   |||  | ||| |||:  :
```

```
Qy     30  KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL    89
           ||:||:: ||  :||||:||   |:: |:|||||::|||||| :|:::||
Db      7  RKVLVANRGEIAIRVFRACTELGIRTVAIYSKEDVGSYHRYKADEAYLVGEGKKPIRAYL    66

Qy     90  DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT   149
           || :||  || ||:||:||||||:|||.|:|.|||.|| ||| :  :|:  ||| ::||
Db     67  DIEGIIEIAKAHDVDAIHPGYGFFSENIQFAKRCREEGIIFFGPNENHLDMFGDKVKARH   126

Qy    150  AAKKAGLPVLAEST-PSKNIDDIVKSAEGQTPIFVKAVAGGGRGMRFVSSPDELRKLA     208
           ||  ||:||: ||  | :|  |:|: ||: ||: |||| |||||||||| :  ||::|
Db    127  AAVNAGIPVIPGSDGPVDGLEDVVAFAEAHGYPIIIKAALGGGGRGMRFVRSKSEVKEAF   186

Qy    209  TEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEI   268
           ||:. ||: ||  .: :||:: :||:|||||||||  | :||||||||.:||||||||||
Db    187  ERAKSEAKAAFGSDEVVVEKLIENPKHIEVQILGDYEGNIVHLYERDCSVQRRHQKVVEV   246

Qy    269  APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLV--DEKGNHVFIEMNPRIQVEHTV   326
           ||:  |. |||:|||  ||: | |:||::||||||||  ||   | |:|:|||||||||:
Db    247  APSVSLSDELRQRICEAAVQLMRSVGYVNAGTVEFLVSGDE---FYFIEVNPRIQVEHTI   303

Qy    327  TEEVTEVDLVKAQMRLAAGATL--KELGL-TQDKIKTHGAALQCRITTEDPNNGFRPDTG   383
           || :  :|:|::|  |  |. |  .| |: .|:. | .| |:| | ||||| |.| ||||
Db    304  TEMITGIDIVQSQILIADGCSLHSHEVGIPKQEDIRINGYAIQSRVTTEDPLNNFMPDTG   363

Qy    384  TITAYRSPGGAGVRLD-GAAQLGGETRAHFDSMLVKMTCRGSDFETAVARAQRALAEFTV   442
           .| |||| ||.||||| |..: |  . || ||:|||  . : |..: |::::| | .| :
Db    364  KIMAYRSGGGFGVRLDAGNGFQGAVITPYYDSLLVKLiSTWALITFEQAARKMLRNLRTFRI   423

Qy    443  SGVATNIGFLRALLREDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKP    502
           .   ||| || |::.    |:  . || ||.| |:. |     :.| |  | .| |||
Db    424  RVFKTNIPFLENVVQHPKFLSGEYDTSFIDTTPELFVFPRRKDRGTKMLTYIGTVTVNGF   483

Qy    503  HGVRPKDVAAPI--DKLPNIKDL-PLPRGSRDRLKLGPAAFARDLREQDALAVTDTTF    558
              |. :|| .|  |||| | |:   : :|.|   |  |  ||::|.|: |.|:|||||
Db    484  PGIGKK--KKPVFDKPRLPKLSEAEPIPAGTKQILDKHGPEGLVRHQEQPRVLLTDTTF   541

Qy    559  RDAHQSLLATRVRSFALKPAAEAVAKITPELLSVEAWGGATYDVAMRFLFEDPWDRLDEL   618
           ||||||||||||| . :  |||..  | |.|:|:|.||||| |||:|||:|||||||::|
Db    542  RDAHQSLLATRVRTVDIVRAAEPSARLLPNLFSLEMWGGATFDVAYRFLKEDPWDRLLKL   601

Qy    619  REAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKRAASSGVDIFRIFDALNDVSQMRPAI   678
           |:| |||.:|||||..|.||| .||| | |.|| ::| :| .:|||||:||.|  | .||
Db    602  RDAFPNVLFQMLLRSANAVGYKNYPDNVIREFVEKSAHAGIHVFRIFDSLNWVKGMTVAI   661

Qy    679  DAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRP   738
           ||| ::   : | |  |:|| .||. | |:|||| | |:||.|:||||||||||||||:|
Db    662  DAVRQSC-KIAEAAICYTGDILDPSRSKYNLDYYKALAKELEQAGAHILAIKDMAGLLKP   720

Qy    739  AAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSL   798
           .|| .|: :| :|.|:|:|:|||||:|.|: ||: :|:|||| ||.|::::||||||||:
Db    721  QAAHVLISGLKETVDIPIYLHTHDTSGNGIYTYAKAIEAGVDIVDVAISSMAGLTSQPSA   780

Qy    799  SAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLS   858
           ..: .| . |.  ..: ::.:.: :||||:|| :| ||||  . .|. |.|.:|||||.|
Db    781  NTLYYALEGTERAPEVDIYGLEQLARYWEDVRKFYQEFESGMNAPHTEVYMHEMPGGQYS   840

Qy    859  NLRAQATALGLADREFLIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFA   918
           ||  ||: :||.||  ::|:.| .||::. |  |||||||||||:||::| ..: .:|||
Db    841  NLQQQAKAVGLGDRWDEVKEMYRRVNDLFGDIVKVTPSSKVVGDMALYMVQNNLTEQDIF   900

Qy    919  ADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDAD   978
           :.   .:.|||| .:| | ||.| || |: |:|  |  |.||:|  |:||: :| .:.
Db    901  ERGETLNFPDSVVELFEGYLGQPHGGFPEKLQRIILLKGRE---PITVRPGELLEPVDFE   956

Qy    979  D-SKERRNSLNR-------LLFPKPTEEFLEHRRFGNTSALLDREFFYGLVEGRET      1027
             :| :  .|:.       | :|  :|||:| :| ||.|:||||  ||:|::|:|
Db    957  QMKKELYDKLGKLGREVIDFDAIAYALYPKVFLEYAETVEKYGDVSVLDPTFLYGMRLGEEI  1016

Qy   1028  LIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAERKADSS  1087
           : : :|.  .|:|:|  |: | .|.  :|. :|||| |.|.||||||: :   || ||.:
Db   1017  EVEIERGKT-LIVKLVSIGQPQADGTRVVYFELNGQPREVVIRDESIKTAVVEHIKADRT  1075

Qy   1088  NKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVPAATK   1146
           |  | ||: .| |||: |.|| .|.|||: :::||||||:||:.:   |:|:: : .
Db   1076  NPNHIAAIMPGTVVKVLVEKGEKVDKGDHLMVTEAMKMETTVQAPFAGVVKDIYKSGDA   1135

Qy   1147  VEGDLIVVVS   1157
           :: |||:| ::
Db   1136  IQAGDLLMELS   1146
```

RESULT 14
Q9HES8

```
ID   Q9HES8          PRELIMINARY;            PRT;   1192 AA.
AC   Q9HES8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase (EC 6.4.1.1).
GN   PYC.
OS   Aspergillus niger.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX   NCBI_TaxID=5061;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N400;
RA   Panneman H., Ruiter G.J.G., Van den Broeck H.C., Visser J.;
RT   "Aspergillus niger pyruvate carboxylase.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- COFACTOR: BIOTIN (BY SIMILARITY).
DR   EMBL; AJ009972; CAC19838.1; -.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; Biotin_carb_C.
DR   InterPro; IPR005481; CPase_L_D2.
DR   InterPro; IPR000891; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; Pyruv_Carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Biotin; Ligase.
SQ   SEQUENCE 1192 AA;  130866 MW;  C7630C689BA23A64 CRC64;
```

```
Query Match      42.8%;   Score 2507.5;   DB 3;   Length 1192;
Best Local Similarity   46.6%;        Pred. No. 1.5e-125;
Matches   545;   Conservative 176;   Mismatches 404;   Indels 45;   Gaps 17;

Qy     19  STHT---SSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASE    74
                    :||:||||||||||:||||||| ||| |:|||||.|:||||||||||
Db     27  SVHTRLRANSAIMQFQKILVANRGEIPIRFRTAHELSLQTVAYSHEDHLSMHRQKADE    86

Qy     75  AVRIGTEG--SPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIG   132
            | :||  .  :||.||.|||| :|| |: || :|:||||||||||| |||.  :||:|.|
Db     87  AYMIGKRGQYTPVGAYLAIDEIVKIALEHGVHLIHPGYGFLSENAEFARKVEQSGMVFVG   146

Qy    133  PTPEVLDLTGDKSRAVTAAKKAGLPVLAEST-PSKNIDDIVKSAEGQTYPIFVKAVAGGG   191
            |||:.::| |||  : .: |  .:|| | .|  . .: |:||  ..:|:||.|||| |||
Db    147  PTPQTIESLGDKVSARQLAIRCDVPVVPGTPGPVERYEVKAFTDTYGFPIIIKAAFGGG   206

Qy    192  GRGMRFVSSPDELRKLATEASREAEAAPGDGSVYVERAVINPOHIEVQLLGDRTGEVVHL   251
            ||||| |:  :||::    |:||::|| | |:| |||| ::|: ||||||||. .::|||
Db    207  GRGMRVVRDQAELRDSFPRATSEARSAFGNGTVFVERFLDRPKHIEVQLLGDNHGNVVHL   266

Qy    252  YERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNH   311
```

```
OS   Listeria monocytogenes.
OC   Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX   NCBI_TaxID=1639;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EGD-e / Serovar 1/2a;
RX   MEDLINE=21537279; PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
DR   EMBL; AL591977; CAC99150.1; -.
DR   ListList; LMO1072; -.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; Pyruv_carbox; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   1146 AA;  128049 MW;  A062F88C9A092B6F CRC64;
```

```
Query Match             42.5%;  Score 2494;  DB 16;  Length 1146;
Best Local Similarity   45.7%;  Pred. No. 7.7e-125;
Matches  525;  Conservative 186;  Mismatches 408;  Indels 30;  Gaps 11;

Qy         30 KKILVANRGEIAVRAFRAALETGAAVIAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL 89
              :| | :||||||||||||||||| |||:||:| ::|   | |||: || ::||  ||| |
Db          5 KKVLVANRGEIAIRVMRACTELKIKTVAIYSEDTGSFHRYKSDEAIVGAGKPIDAYL 64

Qy         90 DIDEIIGAAKKVKADAIYPGVGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT 149
              ||| :|| ||:|:|||||||||||||:||||||:||:||||||||||||:|||||||||
Db         65 DIENIIEIAKESGADAIHPGYGFLSENIEFARRCEQGIIFVGPKSKHLDMFGDKIKAKE 124

Qy        150 AAKKAGLPVLAEST-PSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKLA 208
              ||:||||||:|:|  |||:|||||||| | || |||||| |||||||||:|:| |:| |
Db        125 QALLADIPVIPGSNGPVAGIKEVEEFGEKNGYPLMIKASLGGGGRGMRVVESKEHVKESF 184

Qy        209 TEASREAEAAFGDGSYVVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEI 268
              || :| |||| | || | || :::| || |||||: :| :| |||||:||:|||||||
Db        185 ERASSEAKAAFGNDEVYVEKCVMNPKHIEVQILGDTHGNIVHLFERDCSIQRRHQKVVEV 244

Qy        269 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE 328
              ||| ::|:|:|| ||| ||::|| :|| | || | |||  ||:|| |||:|:||:||| |
Db        245 APCNAITSELRNRICDAAVKLMKNVDYINAGTVEFLV-EGDDFYIEVNPRVQVEHTITE 303

Qy        329 EVTEVDLVKAQMRLAAGATLKELGLT---QDKIKTHGAALQCRITTEDPNNGFRPDTGTI 385
              : |:|:|:|| :|:: ::| | |: |    | ||:|||:|| ||||| :| |:|||| |
Db        304 MITGIDIVQSQLFIADGYALHDQLVAIPKQEDIHIHGSAIQSRITTEDPLNNFMPDTGRV 363

Qy        386 TAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSG 444
              |||||||| | ||| |  ||||||  |||:| ||||||||||||| | ||||| ||| |
Db        364 DTYRSTGGFGVRLDAGNGFQGTVVTPFYDSLLVKLCTWGMTFEQATRKMRNLIEFRIRG 423
```

```
RESULT 15
Q8Y846
ID   Q8Y846        PRELIMINARY;  PRT;  1146 AA.
AC   Q8Y846;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   PycA protein.
GN   PYCA OR LMO1072.
```

Left alignment column:

```
Db   267 :||||||||||||||| |||| ||| :: ||    :                         326
         FERDCSVQRRHQKVVEIAPAKDLPADVRDRILLADAVKLAKSVNYRNAGTAEFLVDQQNRY
Qy   312 VFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITT  371
         :||::|||||||||:|| :|:: ::||    :||||||::||||::||||||||||||||
Db   327 YFIEINPRIQVEHTTEEITGIDIVAAQIOIAAGATLEQLGLTQDRISTRGFAIQCRITT   386
Qy   372 EDPNNGFRPDTGTTTAYRSPGGAGVRLDGAAQLGGE-ITAHFDSMLVKMTCRGSDFETAV  430
         |||: ||||||| |||||| || | ||| || |||  ||||||:| |||||||||| :|
Db   387 EDPSKGFSPDTGKIEVYRSAGGNGVRLDGGNGFAGAIITPHYDSMLVKCTCRGSTYEIAR  446
Qy   431 ARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRI  490
         |:: |:||||:  | :| | :  || || ::  || ||||| :::|   |:  :|::: :
Db   447 RKVVRALVEFRIRGVKTNIPFLTSLLSHPVFVFVDGTCWTTFIDDTPELFALLVGSNRAQKL  506
Qy   491 LDYLADVTVN----KPHGVRPK---DVAAPI--DKLPNIKDLPLP--RGSRDRLKQLGPA  539
         | |:|  |    ||  |:|:   | ::  | |:| ||:| | | :  :|::|:| |:|
Db   507 LAYLGDVAVNGSSIKGQIGEPKLKGDIIKPVLHDAAGKPLDVSVPATKGWKQILDSEGPE  566
Qy   540 AFARDLREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGAT  599
         ||| :| : ||| : || ::||||||||||:|| ||:||||:||:| ||||| |||||||
Db   567 AFARAVRANKGCLIMDTTWRDAHQSLLATRVRTIDLLNIAHETSHALANAYSLECWGGAT  626
Qy   600 YDVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGTPYPDSVCRAFVKEAASSGV  659
         :|||||| |||||||| |:|   |:|: |:|||| |:|:|||| ||| |||: || |
Db   627 FDVAMRFLYEDPWDRLRKLRKAVPNIPFQMLLRGANGVAYSSLPDNAIYHFCKQAKKCGV  686
Qy   660 DIFRIFDALNDVSQMRPAIDAVLENTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEI  719
         ||||:||||||| :| | |||| |||  |  ||  || |:: :||| ||:| ||:: :
Db   687 DIFRVFDALNDVDQLEVGIKAV-HAAEGVVEATICYSGDMLNPSKK-YNLPYYLDVDKV  744
Qy   720 VKSGAHILAIKDMAGLLRPAAVTKIVTALRREF-DLPVHVTHDTAGGQLATYFAAAQAG  778
         |: || ||:|| |::  :|  |:| |  || |  ::|||:|||: :|: || ::||||
Db   745 VQFKPHVLGIKDMAGVLKPQAARLLIGSIRERYPDLPIHVTHDSAGTGVASMIACAQAG  804
Qy   779 ADAVDGASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFES  838
         ||||:  | |  ||| |:||  | |: ||:| ||  :|| | | ||::|:| |||| |:|:
Db   805 ADAVDAATDSLSGWTSQPSIGAILASLSEGTEHDPGLNSAQVRALDTYWAQLRLLYSPFEA  864
Qy   839 GTPGPTGRVYRHEIPGGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSK  898
         ||:| |||:| |: || ||:  |:|| | | || :|  ::::  |:  |:| | ||| |
Db   865 GLTGPDPEVYEHEIPGGQLTNLIFQASQLGLGQQWAETKKAYESANDLLGDVVKVTPTSK  924
Qy   899 VVGDLALHVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPGGWPEPLRTRALEGRS  958
         |||| | ||| ||  :|: :|| | |  :|:||| | |: ||| || |:|| || :
Db   925 VVGDLAQFMVSNKLTAEDVIARAGELDFPGSVLEFLEGLMGQPYGGFPEPLRSRALRDRR  984
Qy   959 E-GKAP--------LTEVPEEQAHLDADDSKERRNSLNRLLFPKPTEEFLEHRRR 1005
         :  | |         | :| |: :  :| |   ::::||: || :||:: | :
Db   985 KLDKRPGLYLEPLDLAKIKSQIRENYGAATEYDVA------SYAMYPKVFEDYKKFVAK  1037
Qy  1006 FGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISE-PDDKGMRNVVANVNGQI 1064
         :    | || |:|  :|   ||:   |  | |:|  :| ::   |||:| :|| || |:
Db  1038 FGDLSVLPTRYFLAKPEIGEEFHVELEKGKV-LILKLLAIGPLSEQTGQREVFYEVNGEV  1096
Qy  1065 RPMRVRDRSVESVTATAEKADSSNKGHVAAPFAGVVT-VTVAEGDEVKAGDAVAIIEAMK 1123
         ||::| |  | | ||  : |::     | | | ||    |: | || | |||||::| :|
Db  1097 RQVSVDDKKASVENTARPKAELGDSSQVGAPMSGVVVEIRVHDGLEVKKGDPIAVLSAMK 1156
Qy  1124 MEAITASVDGKIERVVVPAATKVEGGDLI 1153
         ||: |  :|| |  ::  ::  | || |
Db  1157 MEMVISAPHSGKVSSLLVKEGDSVDGQDLV 1186
```

```
Qy   445  VATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKPHG  504
              |||||| :: ::||   |||    | | |  |    || :::  | |:  :|||
Db   424  VKTNIPFLLNVVRHPDFASGNYNTSFIDTTPELFKFPHIRDRGIKTLRYIGNVTVNGFPG  483

Qy   505  VRPKDVAAPIDKLPNIKDLP----LPRGSRDRLKQLGPAAFARDLREQDALAVIDTTFRD  560
              ::: |:  |: |  |:::      |  :| | :||  || | :::: :||||| ||
Db   484  IKHRD--KPYYAEPRLPKIPYGSQISPGTKQILDAKGPEGVVDWVKKQEVLLTDTTLRD  541

Qy   561  AHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGANYDVAMRFLFEDPWDRLDELRE  620
              ||||||||||||| : |||: : |  |: :| |||||| ::: |||  ||| | ||
Db   542  AHQSLLATRVRSKDIFQVADAMAHLLPNMESFEMWGGATFDVAYRFLNEDPWVRLETLRK  601

Qy   621  AMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDA  680
              :||| |||||| ||||| |  |||:|    | | ||: ||: :| |  :  ::  :||
Db   602  QIPNVMFQMLLRGANAVGYKNYPDNVIREFVKQSAQSGVDVFRVFDSLNWIKGMEVSIDA  661

Qy   681  VLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAA  740
              |  | : |:||    | |::     |   ||: || ::  | ||| |||||||||||
Db   662  VREAG--KVVEAAICYTGDIDDTRTKYTIDYYKDMAKELVAQGTHILGIKDMAGLLKPQA  720

Qy   741  VTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSA  800
              : |: |:  :|:|   |  ||  ||| ||:   |  | | : :  |: ||::|::: |
Db   721  AYRLIGELKDTVDVPIHLHTHDTSGNGIYTYAAAVSAGVDIVDVASSAMSGATSQPSMTG  780

Qy   801  IVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSNL  860
              : : | |   || | |:: :   |||  | : |:| : ||| ||||||  ||||  ::|
Db   781  LYYGLVNGNRQTNLDAQNSQIINHYWEDVRHYYKDFDNALNSPQTEVYIHEMPGGQYTNL  840

Qy   861  RAQATALGLADRFLIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAAD  920
              || |  ||  ||   :|||   ::  |   ||||| ||||||:|||||:|:||||  |
Db   841  QQQAIAVGLGDRWDEVKEMTVVNQMFGDIVKVTPSSKVVGDLALFMVQNELSEDDVYEK  900

Qy   921  PQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDADDS  980
              |  | ||||||    ||| | |  ||:     ||| |:  |||   :::   :|  : :
Db   901  GDTIDFPDSVIEFFMGEIGQPYGGFPEKLQKLVLKGRT---PLTDRPGALMEPVNFVDV  956

Qy   981  K-------ERRNSLNRLLFPKPTEEFLEHRRFGNTSALDDREFFYGLVEGRETLI  1029
          |        :: :||:||    :: :|   ||: :| || :: |: : :
Db   957  KAELKEKMGYEPTEKDVISYILYPKVFLDYQDMINKYGDVTVLDTPTFYKGMRLGETIEV  1016

Qy  1030  RLPDVRTPILVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNK  1089
              ||:| |:|:: ::|| ::| |:  :: :||:  : | |:| : | ||::: :  :::
Db  1017  ELEKGKI-LLIKLNSIGEPIADGTRVIYFELNGQPREINIQDMNVQSTVIARRKIDTTNP  1075

Qy  1090  GHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVVPAATKVE  1148
              ||  ||:|  :| |||| |:|| ||:|::|| |: || :|| |||| :| |:|  | |:
Db  1076  EHVGATMIGSVIQVVVKKGDSVKKGDPLLITEAMKMETTIQAPFDGEVSSIYVSDGDTIE  1135

Qy  1149  GGDLIVVVS  1157
              |||:|:||:
Db  1136  SGDLLIEVN  1144
```

```
              Copyright   GenCore version 5.1.6
              Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:     September 24, 2003, 15:51:28 ; Search time 88.3413 Seconds
                                           (without alignments)
                                           3330.038 Million cell updates/sec

Title:      US-09-974-973A-19
Perfect score:  5788
Sequence:   1 MSTHYSSTLPAFKKILVANR.........RVVVPAATKVEGGDLIVVVS 1140

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5
Searched:   830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters:        830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  SPTREMBL_23:*
             1: sp_archea:*
             2: sp_bacteria:*
             3: sp_fungi:*
             4: sp_human:*
             5: sp_invertebrate:*
             6: sp_mammal:*
             7: sp_mhc:*
             8: sp_organelle:*
             9: sp_phage:*
            10: sp_plant:*
            11: sp_rodent:*
            12: sp_virus:*
            13: sp_vertebrate:*
            14: sp_unclassified:*
            15: sp_rvirus:*
            16: sp_bacteriap:*
            17: sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5788	100.0	1140	16	O54587	O54587 corynebacte
2	5759	99.5	1140	16	Q8L2G4	Q8l2g4 corynebacte
3	5317.5	91.9	1168	16	Q8FRQ0	Q8frq0 corynebacte
4	5305.5	91.7	1139	2	Q8RQL2	Q8rql2 corynebacte
5	3799	65.6	1124	16	Q9RK64	Q9rk64 streptomyce
6	3783.5	65.4	1127	2	Q9F843	Q9f843 mycobacteri
7	3712.5	64.1	1127	16	P95127	P95127 mycobacteri
8	3483	60.2	1124	2	Q50450	Q50450 mycobacteri
9	2552	44.1	1178	13	Q8JHF6	Q8jhf6 gallus gall
10	2525.5	43.6	1150	16	Q9K9M0	Q9k9m0 bacillus ha
11	2524	43.6	1148	16	Q9KWU4	Q9kwu4 bacillus su
12	2519.5	43.5	1180	13	Q9DDT1	Q9ddt1 brachydanio
13	2509.5	43.4	1192	3	Q9HES8	Q9hes8 aspergillus
14	2505	43.3	1147	2	P94448	P94448 bacillus st
15	2493	43.1	1146	16	Q8Y846	Q8y846 listeria mo
16	2491.5	43.0	1144	16	Q97FR7	Q97fr7 clostridium
17	2490	43.0	1146	16	Q92CW1	Q92cw1 listeria in
18	2489	43.0	1150	16	Q8NX69	Q8nx69 staphylococ
19	2488	43.0	1150	16	Q99UY8	Q99uy8 staphylococ
20	2487.5	43.0	1193	3	O93918	O93918 aspergillus
21	2477.5	42.8	1158	16	Q8FYT0	Q8fyt0 brucella su
22	2473.5	42.7	1174	16	Q8UBX3	Q8ubx3 agrobacteri
23	2470.5	42.7	1152	16	Q98F27	Q98f27 rhizobium l
24	2464.5	42.6	1158	16	Q8YJ20	Q8yj20 brucella me
25	2464	42.6	1153	16	Q8CPM3	Q8cpm3 staphylococ
26	2464	42.6	1175	5	O17732	O17732 caenorhabdi
27	2462.5	42.5	1179	13	Q8AYN3	Q8ayn3 pagrus majo
28	2443	42.2	1185	3	Q9UUE1	Q9uue1 schizosacch
29	2441	42.2	1181	5	Q9XZ00	Q9xz00 drosophila
30	2437.5	42.1	1175	16	Q8XlT3	Q8xlt3 pichia angu
31	2435	42.1	1152	16	Q92L13	Q92l13 rhizobium m
32	2435	42.1	1185	3	P78822	P78822 schizosacch
33	2435	42.1	1197	16	Q8MKW5	Q8mkw5 drosophila
34	2419	41.8	1195	16	Q16921	Q16921 aedes aegyp
35	2416	41.7	1147	16	Q8ER83	Q8er83 oceanobacil
36	2406	41.6	1137	16	Q9CHQ7	Q9chq7 lactococcus
37	2402	41.5	1137	2	Q9RAT6	Q9rat6 lactococcus
38	2400	41.5	1154	2	O59740	O59740 rhizobium e
39	2155.5	37.2	1029	13	Q8AXQ6	Q8axq6 pagrus majo
40	2115.5	36.5	984	2	Q9XBJ1	Q9xbj1 bacillus ce
41	2017	34.8	935	11	Q62043	Q62043 mus musculu
42	1930.5	33.4	920	2	Q9KWU5	Q9kwu5 bacillus su
43	1658	28.6	828	11	Q8BP54	Q8bp54 mus musculu
44	1100	19.0	525	5	Q8MKW4	Q8mkw4 drosophila
45	1099	19.0	541	5	Q8MKW3	Q8mkw3 drosophila

ALIGNMENTS

```
RESULT 1
O54587
ID   O54587       PRELIMINARY;      PRT;   1140 AA.
AC   O54587;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase (EC 6.4.1.1).
GN   PYC OR CGL0689.
OS   Corynebacterium glutamicum (Brevibacterium flavum).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=1718;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=21253;
RA   Koffas M.A.G., Ramamoorthi R., Pine W.A., Sinskey A.J.,
RA   Stephanopoulos G.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 13032;
RA   Peters-Wendisch P.G., Kreutzer C., Kalinowski J., Patek M., Sahm H.,
RA   Eikmanns B.J.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA   Nakagawa S.;
RT   "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- COFACTOR: BIOTIN (BY SIMILARITY).
DR   EMBL; AF038548; AAB92588.1; -.
DR   EMBL; Y09548; CAA70739.1; -.
DR   EMBL; AP005276; BAB98082.1; -.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005481; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
```

```
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Biotin; Ligase; Pyruvate; Complete proteome.
SQ   SEQUENCE 1140 AA;  123102 MW;  2A6D4B4ED2FEB531 CRC64;

Query Match             100.0%;  Score 5788;  DB 16;  Length 1140;
Best Local Similarity   100.0%;  Pred. No. 1.9e-302;
Matches 1140;   Conservative    0;   Mismatches    0;   Indels    0;   Gaps    0;

Qy     1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60

Qy    61 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120

Qy   121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180

Qy   181 VASPDELRKLATEASREAEAAFGDGAVVVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 VASPDELRKLATEASREAEAAFGDGAVVVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240

Qy   241 LQRRHQKVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 LQRRHQKVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 300

Qy   301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360

Qy   361 FRPDTGTITAYRSPGGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 FRPDTGTITAYRSPGGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420

Qy   421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRIILDYLADV 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRIILDYLADV 480

Qy   481 TVNKPHGVRPKDVAAPIDKLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TVNKPHGVRPKDVAAPIDKLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 540

Qy   541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATVDVAMRFLFEDPWDRLDE 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATVDVAMRFLFEDPWDRLDE 600

Qy   601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 660

Qy   661 IDAVLETNTAVADEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 IDAVLETNTAVADEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR 720

Qy   721 PAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 PAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 780

Qy   781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQL 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQL 840

Qy   841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900

Qy   901 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA 960

Qy   961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020

Qy  1021 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080

Qy  1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS 1140


RESULT 2
Q8L2G4
ID   Q8L2G4          PRELIMINARY;           PRT;   1140 AA.
AC   Q8L2G4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase.
GN   PYC.
OS   Corynebacterium crenatum.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=168810;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CD945;
RA   Wang J., Ding J., Liu Y.;
RT   "Cloning and Expression of Pyruvate Carboxylase Gene in
RT   Corynebacterium crenatum CD945.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF503915; AAM27458.1; -.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
SQ   SEQUENCE 1140 AA;  123126 MW;  FFA90BB7644C910E CRC64;

Query Match              99.5%;  Score 5759;  DB 2;  Length 1140;
Best Local Similarity    99.3%;  Pred. No. 6.9e-301;
Matches 1132;   Conservative    5;   Mismatches    3;   Indels    0;   Gaps    0;

Qy     1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60
```

```
RESULT 3
Q8FRQ0
ID   Q8FRQ0       PRELIMINARY;      PRT;   1168 AA.
AC   Q8FRQ0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase (EC 6.4.1.1).
GN   PYC OR CE0709.
OS   Corynebacterium efficiens.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=152794;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA   Kawarabayasi Y., Yamazaki Y., Hino Y., Kikuchi H., Nakamura Y.,
RA   Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
RA   Usuda Y., Sugimoto S.;
RT   "The entire genomic sequence of Corynebacterium efficiens YS-314.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP005216; BAC17519.1; -.
KW   Ligase; Complete proteome.
SQ   SEQUENCE 1168 AA;  126245 MW;  A5D5A4D4DF285F8 CRC64;

Query Match             91.9%;  Score 5317.5;  DB 16;  Length 1168;
Best Local Similarity   91.2%;  Pred. No. 3.8e-277;
Matches 1038;           Conservative 43;  Mismatches 56;  Indels 1;  Gaps 1;
```

```
Qy      3 THTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIG 62
          | ||||| ||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db     32 TTTPSTLPAFKKILVANRGEIAVRAFRAAYETGAATVAIYPREDRGSFHRSFASEAVRIG 91

Qy     63 TEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLD 122
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     92 TEGSPVKAYLDIDEIINAAKKVKADAVYPGYGFLSENAQLARECAENGITFIGPTPEVLD 151

Qy    123 LTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYIPFVKAVAGGGRGMRFVA 182
          |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db    152 LTGDKSKAVSAAKKAGLPVIAESTPSTDIDEIVKSAEGQTYIPFVKAVAGGGRGMRFVE 211

Qy    183 SPDELRKLATEASREARAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQ 242
          |:::||||||||||||||||||||||||||||||||||||:||:|||||||||||||||
Db    212 KPEDLRELAREASREAEAAFGDGSVYVERAVIKPQHIEVQILGDHTGDVIHLYERDCSLQ 271

Qy    243 RRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPR 302
          |||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db    272 RRHQKVVEIAPAQHLDPELRDRICADAVKFCKSIGYQGAGTVEFLVDEAGNHVFIEMNPR 331

Qy    303 IQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFR 362
          |||||||||||:|||||||||||||||||||||||||||||||||||||||||||||:||
Db    332 IQVEHTVTEEVTSVDLVKAQMHLAAGATLKELGLTQDKITTHGAALQCRITTEDPSNNFR 391

Qy    363 PDTGTTAYRSPGGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRALAE 422
          ||||||||||||||||||||||||||||||||||:|||||||||||||:|||:|||||
Db    392 PDTGVTTAYRSPGGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVSRAQRALAE 451

Qy    423 FTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTV 482
          | |||||||||||||||||||||||||||||:|| ||||||||||||||||||:|||||
Db    452 FNVSGVATNIGFLRALLREEDFTKRRIDTGFIGSHQHLLQAPPADDEQGRILEYLADVTV 511

Qy    483 NKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTTFR 542
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db    512 NKPHGERP-ETARPIEKLPEVENIPLPRGSRDRLKQLGPEGFARDLREQDALAVTDTTFR 570

Qy    543 DAHQSLLATRVRSFALKPAAEVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELR 602
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    571 DAHQSLLATRVRSFALTPAARAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWARLDELR 630

Qy    603 EAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAID 662
          ||||||||||||||||||||||||||||||||||:| ||||||||||||||||||||||
Db    631 EAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVQEAAKSGVDIFRIFDALNDISQMRPAID 690
```

The second column (continuing alignments):

```
Qy    61 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 IGTEGSPVKAYLDIDEIINAAKKVKADAVYPGYGFLSENAQLARECAENGITFIGPTPEV 120

Qy   121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYIPFVKAVAGGGRGMRF 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 LDLTGDKSRAVTAAKKAGLPVIAESTPSKNIDDIVKSAEGQTYIPFVKAVAGGGRGMRF 180

Qy   181 VASPDELRKLATEASREARAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240
         |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 VSSPDELRKLATEASREARAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCS 240

Qy   241 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 300

Qy   301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360

Qy   361 FRPDTGTITAYRSPGGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 FRPDTGTITAYRSPGGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420

Qy   421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADV 480

Qy   481 TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db   481 TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 540

Qy   541 FRDAHQSLLATRVRSFALKPAAEVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 FRDAHQSLLATRVRSFALKPAAEVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600

Qy   601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAATSGVDIFRIFDALNDVSQMRPA 660

Qy   661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 IDAVLETNIAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR 720

Qy   721 PAAVTKLVTALRREEDLPVIVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 780
         ||| ||||||||||||| |||||||||||||||| ||||||||||||||||||||||||||
Db   721 PAAATKLVTALRREEDLPVIVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 780

Qy   781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840

Qy   841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900

Qy   901 AADPQKYDIPDSVIAFLRGELGNPPGGWEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 AADPQKYDIPDSVIAFLRGELGNPPGGWEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA 960

Qy   961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREEFYGLVEGRETLIRLPDVRTP 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREEFYGLVEGRETLIRLPDVRTP 1020

Qy  1021 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080

Qy  1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db  1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVIS 1140
```

```
Query Match              91.7%;  Score 5305.5;  DB 2;  Length 1139;
Best Local Similarity    91.0%;  Pred. No. 1.6e-276;
Matches 1037;  Conservative 43;  Mismatches 59;  Indels  1;  Gaps  1;

Qy    1 MSTHTSSTLPAFKKILVANGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60
          | ||||||||||||||||||||| :| |||||||||||||||||||||||||||||||
Db    1 MVTTTPSTLPAFKKILVANGEIAVRAFRAAYETGAATVAIYPREDRGSFHRSFASEAVR 60

Qy   61 IGTEGSPVKAYLDIDEIIGAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120
          |||||||||||||||||| ||||||||| ||||||||||||||||||||||||||||||
Db   61 IGTEGSPVKAYLDIDEIINAAKKVKADAVYPGYGFLSENAQLARECAENGITFIGPTPEV 120

Qy  121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEQTYPIFVKAVAGGGRGMRF 180
          |||||||| ||| ||||||||||||||| ||||||||||||||||||||||||||||||
Db  121 LDLTGDKSAVSAAKKAGLPVLAESTPSTDIDEIVKSAEQTYPIFVKAVAGGGRGMRF 180

Qy  181 VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240
          | ::|||| ||||||||||||||| ||||||||| ||||||||||:: ||||||||||||
Db  181 VEKPEDLRELAREASREAEAAFGDGSVYVERAVIKPQHIEVQILGDHTGDVIHLYERDCS 240

Qy  241 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 300
          ||||||||||||||||||||||||||||||||| |||||||||||||||| ||||||||
Db  241 LQRRIQKVVEIAPAQHLDPELRDRICADAVKFCKSIGYQGAGTVEFLVDEAGNHVFIEMN 300

Qy  301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATIKELGLTQDKIKTHGAALQCRITTEDPNNG 360
          |||||||||||||| ||||||||| |||||:|||||||||| |||||||||||||||:
Db  301 PRIQVEHTVTEEVTSVDLVKAQMHLAAGATLKELGLTQDKITTHGAALQCRITTEDPSNN 360

Qy  361 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
          |||| || ||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db  361 FRPDFGVITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVSRAQRAL 420

Qy  421 AEFTVSGVATNIGFLRALLREEDTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV 480
          |||:|||||||||||||||||||| :| ||||| :|:|||||||||||||||:||||||
Db  421 AEFNVSGVATNIGFLRALLREEDTKRRIDTGFIGSHQHLLQAPPADDEQGRILEYLADV 480

Qy  481 TVNKPHGVRPKDVAAPIDKLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 540
          |||||||  ||| | |:| || ||||||||||||| :||||||||||||||||
Db  481 TVNKPHGERP-ETARPIEKLPEVENIPLPRGSRDRLKQLGPEGFARDLREQDALAVTDTT 539

Qy  541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600
          ||||||||||||||||| |||| |||||||||||||||||||||||||||||||| ||||
Db  540 FRDAHQSLLATRVRSFALTPAARAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWARLDE 599

Qy  601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 660
          ||||||||||||||||||||||||||||||||||:|||  |||||||||||::||  ||||
Db  600 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVQEAAKSGVDIFRIFDALNDISQMRPA 659

Qy  661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYIKMAEEIVKSGAHILAIKDMAGLLR 720
          |||||| |:|||||||||||||| | ||||||||: |||:||:|||||||||||||||
Db  660 IDAVLETGTSVAEVAMAYSGDLSNPGEKLYTLDYLNLAEQIVDSGAHILAIKDMAGLLR 719

Qy  721 PAAVTKLVTALRREFDLPVHVTHTDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 780
          || | ||| ||||||||||||||||||||||||| |||  |||||| |||||||||||||
Db  720 RAAAPKLVIALRREFDLPVHVTHTDTAGGQLATYLAAANAGADAVDAASAPLSGTTSQPS 779

Qy  781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQL 840
          :||:||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  780 MSALVAAFAHTRRDTGLNLQAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQL 839

Qy  841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900
          ||||||| |||||||||||||:||||||||||||||||||||||||||||||||: ||
Db  840 SNLRAQAVALGLADRFELIEDYYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVSPEDF 899

Qy  901 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDA 960
          ||||||||||||||||||||:|||||||||||||||||||:| ||||:|  |||||||
Db  900 AADPQKYDIPDSVIAFLRGELGTPPGGWPEPLRTRALEGRSQGKAPLAEIPAEQAHLDS 959

Qy  961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020
          :||  |||| | |||||||||||||||||||||||||||||||:|||| ||| | :||
Db  960 DDSAERRGTLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLKEGREELIRLTGVSTP 1019
```

```
Qy  663 AVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYILKMAEEIVKSGAHILAIKDMAGLLRPA 722
          |||||:|:||||||||||||||:| ||||||||| :|| ||:|||||||||||||||||
Db  691 AVLETGTSVAEVAMAYSGDLSNPGEKLYTLDYILNLAEQIVDSGAHILAIKDMAGLLRPA 750

Qy  723 AVTKLVTALRREFDLPVHVTHTDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLS 782
          | | ||| ||||||||||||||||||||||| |||  |||||| |||||||||||||
Db  751 AATKLVTALRREFDLPVHVTHTDTAGGQLATYLAAANAGADAVDAASAPLSGTTSQPSMS 810

Qy  783 AIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQLSN 842
          |:|||||||||||| ||||||||||||||||||||||||||||||||||||||||| |
Db  811 ALVAAFAHTRRDTGLNLQAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQLSN 870

Qy  843 LRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADDAA 902
          ||||| |||||||||||||:||||||||||||||||||||||||||||||||| ||| ||
Db  871 LRAQAVALGLADRFELIEDYYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVSPEDFAA 930

Qy  903 DPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDADD 962
          ||||||||||||||||||:|||||||||||||||||||:| |||| :|  |||||| ||
Db  931 DPQKYDIPDSVIAFLRGELGTPPGGWPEPLRTRALEGRSQGKAPLAEIPAEQAHLDSDD 990

Qy  963 SKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLL 1022
          | |||| | |||||||||||||||||||||||||||||||:||||  || | :|| :
Db  991 SAERRGTLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLKEGREELIRLTGVSTPMV 1050

Qy 1023 VRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAGV 1082
          |||| ||||||||||| | ||:||||:|:||||:||||||||| ||: |||||||| |||
Db 1051 VRLDAVSEPDDKGMRNVVVNVNGQIRPIKVRDRSVESVTATAEKADATNKGHVAAPFAGV 1110

Qy 1083 VTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS 1140
          |||||||||||:|||| ||||||||||||| | ||| |||||||||||||||||||||
Db 1111 VTVTVAEGDEIKAGDAVAIIEAMKMEATITAPVDGVIDRVVVPAATKVEGGDLIVVVS 1168
```

```
RESULT 4
Q8RQL2
ID   Q8RQL2           PRELIMINARY;      PRT;   1139 AA.
AC   Q8RQL2;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase.
GN   PYC.
OS   Corynebacterium efficiens.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=152794;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Akiyoshi N., Nonaka G., Kimura E., Kawahara Y., Sugimoto S.;
RT   "Corynebacterium efficiens Pyruvate carboxylase (pyc) gene, complete
RT   CDS.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- COFACTOR: BIOTIN (BY SIMILARITY).
DR   EMBL; AB083299; BAB88903.1; -.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; Pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Biotin.
SQ   SEQUENCE   1139 AA;  123068 MW;  BA7023134519FAAA CRC64;
```

```
Qy  1021 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEBKADSSNKGHVAAPFA 1080
              ::||||||:||:|||||||||| |||||||| ||||||||||||||  ||:|||||||||
Db  1020 MVVRLDAVSEPDDKGMRNVVNVNVNQIRPIKVRDRSVESVTATAEKADATNKGHVAAPFA 1079

Qy  1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS 1140
              ||||||||||||:|| |||||||||||||||| ||| ||||||||||||||||||||||
Db  1080 GVVTVTVAEGDEIKACDAVAIIEAMKMEATITAPVDGVIDRVVVPAATKVEGGDLIVVVS 1139
```

RESULT 5
Q9RK64
ID   Q9RK64   PRELIMINARY;   PRT;   1124 AA.
AC   Q9RK64;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase.
GN   SCO0546 OR SCF11.26c.
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Seeger K.J., Harris D.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=97000351; PubMed=8843436;
RA   Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA   Kinashi H., Hopwood D.A.;
RT   "A set of ordered cosmids and a detailed genetic and physical map for
RT   the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL   Mol. Microbiol. 21:77-96(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=21996410; PubMed=12000953;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
DR   EMBL; AL939105; CAB59603.1; -.
     HSSP; P24182; 1BNC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.

DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   1124 AA; 121071 MW; 1FEF4C4643954D31 CRC64;

Query Match   65.6%; Score 3799; DB 16; Length 1124;
Best Local Similarity   65.2%; Pred. No. 1.4e-195;
Matches   736; Conservative 154; Mismatches 228; Indels 10; Gaps 6;

```
Qy    12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY  71
            |::||||||||||:|||||| |:||| |||:| |:|||| |:|:: :||:|:| :::|
Db     2 FRKVLVANRGEIAIRAFRAGYELGARTVAVFPHEDRNSLHRLKADEAYEIGEQGHPVRAY  61

Qy    72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 131
           | :::|| |:|| |||||:|||||||||| :|||||| :|||||||| :|:|:|| |:|||||
Db    62 LSVEEIVRAARRAGADAVYPGYGFLSENPELARACEEAGITFVGPSARILELTGNKARAV 121

Qy   132 TAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKLA 191
           :|| ||:||| :|: ||:|:|:||:||::||:|:||||||||||||||||:   :|:|:|:
Db   122 AAAREAGVPVLGSSAPSTDVDELVRAADDVGFPVFVKAVAGGGGRGMRVEEPAQLREAI 181

Qy   192 TEASREAEAAFGDGAVYYERAVINPQHIEVQILGDHTGEVHLYERDCSLQRRHQKVVEI 251
           |||||||:|:|||::|: :|||::|:||||||:||: :|:|:|:||||:|||||||::|:|
Db   182 EAASREAASAFGDSTVFLEKAVVEPRHIEVQILADGEGDVIHLFERDCSVQRRHQKVIEL 241

Qy   252 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE 311
           |||:.||||:|:||||||||:| |::|| :||||||||||::||||||||||||||||||||
Db   242 APAPNLDPALRERICADAVNFARQIGYRNAGTVEFLVDRDGNHVFIEMNPRIQVEHTVTE 301

Qy   312 EVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAY 371
           |||:||||::|:|:|||:||::||||::::: ::|||||||||||||:||||||||::||||
Db   302 EVTDVDLVQSQLRIAAGQTLADLGLAQENITLRGAALQCRITTEDPANGFRPDTGQISAY 361

Qy   372 RSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 430
           ||||:.|:|||| :: :||:|||||||||||:|||:|||:||:||||:|:||||  |||
Db   362 RSPGGGSGIRLDGGTTHAGTEISAHFDSMLVKLSCRGRDFTTAVNRARAVAEFRIRGVAT 421

Qy   431 NIGFLRALLREEDFTSKKRIATGFIADHPHLLQAPPADDEQGRIIDYLADVTVNKPHGYRP 490
           || ||:|:| ::.|: :|:::|:|||:::|||| | ::|:| :|:||||||||||||||: ||
Db   422 NIPFLQAVLDDPFDQAGRVTTSFIEQRPHLLFARHSADRGTKLLTYLADVTVNKPHGERP 481

Qy   491 KDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVIDTTFRDAHQSLLA 550
           .::.:|:.|||||  :.  | ||:| | :||| || |::|| |:|||||||||||||||||||
Db   482 -EIVDPLITKLPTASAGEPPAGSRQLLAELGPEGFARRLRESSTIGVIDTTFRDAHQSLLA 540

Qy   551 TRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELEREAMPNVNI 610
           |||:.:::|:|||| ||:|||||||||| |||||||||||:|:||:|||:|:|||| ||||:::|:||:
Db   541 TRVRTKDMLAVAPVVARTLPQLLSLECWGGATYDVALRFLAEDPWERLAALREAVPNLCL 600

Qy   611 QMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETNTA 670
           |||||||||||||||||:||::|||:||::|:|||||||||||||||:|||||:||::::|
Db   601 QMLLRGRNTVGYTPYPTEVTDAFVQEAAATGIDIFRIFDALNDVEQMRPAIEAVRQPGSA 660

Qy   671 VAEVAMAYSGDLSDPNEKLYTLDYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLVTA 730
           |||||:.|:.|||||:|:|||||||::|:|:||::|||:||:|||||||||||  ::||:|
Db   661 VAEVALCYTADLSDPSERLYTLDYLRLAEQIVNAGAHVLAVKDMAGLLRPAAAATLVSA 720

Qy   731 LRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFAH 790
           |||||||||||:|||||:||||||::||:||:|||||||||||::::::|||||||||||::||:
Db   721 LRREFDLPVHLHTHDTVGGQLATYLAAIQAGADAVDGAVASMAGTTSQPSLSAIVAATDH 780

Qy   791 TRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSNLRAQATAL 850
           | | |||.|:||.|||||||:||::|| |||:| |.||||||:|||||||||||:||:||
Db   781 TERPTGLDLQAVGDLEPYWESVRKVYAPFEAGLASPTGRVYHHEIPGGQLSNLRTQAVAL 840

Qy   851 GLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAADPQKYDIP 910
           || ||||:||: ||| :: |||  ||||||||||||||||||||| |||| ||::::|||
Db   841 GLGDRFEDIEAMYAAADRMLGRLVKVTPSSKVVGDLALHLVGAGVSPADFEQDPDRFDIP 900

Qy   911 DSVIAFLRGELGNPPGGWPEPLRTRALEGRSGKAPLTEVPEEQAHLDADDSKERRNSL 970
           |||:.|||||||.|||||||||.:::|||:::| |.:|:.|:::.:.|:|.||::::|
Db   901 DSVVGFLRGELGTPPGGWPEPFRSKALRGRAEAR-PLAELSEDDRDGL---GKDRRATL 955

Qy   971 NRLLFKPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRL-PDVRTPLLVRLDAIS 1029
```

```
Db   956  NRLLFPGPAREFDTHRASYGDTSILDSKDFFYGLRPGKEYTVDLDPGVR--LLIELQAVG 1013
          ||||| ||      ||||||  :  ||||| |||| : ||  | ||  |||:|
Qy  1030  EPDDKGMRNVANVNGQIRPMRVRDRSVESVTATEAKADSSNKGHVAAPFAGVVTVTVAE 1089
          |:::: ||| ::|||| ||  |  ||||||||||||||||||| |
Db  1014  DADERGMRTVMSSLNGQLRPIQVRDRSAATDVPTEKADRANPGHVAAPFAGVVTLAVAE 1073
Qy  1090  GDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVPAATKVEGGDLIV 1137
          ||||| || ||||||||||||  :  ||| :||
Db  1074  GDEVEAGATVATIIEAMKMEASITAPKSGTVTRLAINRIQQVEGGDLLV 1121

RESULT 6
Q9F843
ID   Q9F843
AC   Q9F843;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase (EC 6.4.1.1).
GN   PYC.
OS   Mycobacterium smegmatis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1772;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20374587; PubMed=10913817;
RA   Mukhopadhyay B., Purwantini E.;
RT   "Pyruvate carboxylase from Mycobacterium smegmatis: stabilization,
RT   rapid purification, molecular and biochemical characterization and
RT   regulation at the cellular level.";
RL   Biochim. Biophys. Acta 1475:191-206(2000).
CC   -!- COFACTOR: BIOTIN (BY SIMILARITY).
DR   EMBL; AF262949; AAG30411.1; -.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR000891; CPase_L_N.
DR   InterPro; IPR003379; HMGL-like.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   InterPro; IPR000634; S/T_dehydrtse.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; Pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
DR   PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW   Biotin; Ligase.
SQ   SEQUENCE  1127 AA;  120765 MW;  18C132C4842SC67B CRC64;

Query Match     65.48%;  Score 3783.5;  DB 2;  Length 1127;
Best Local Similarity   66.2%;  Pred. No. 9.7e-195;
Matches  747;  Conservative 141;  Mismatches 234;  Indels 7;  Gaps 6;

Qy    14  KILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLD 73
          |:||||||||| ||||| || |||:|||   ||||   : |||||
Db     4  KVLVANRGEIAIRAFRAAYEMGIATVAVYPYEDRNSLHRLRLKADESYQICEVGHPVRAYLS 63
Qy    74  IDELIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVTA 133
          |:|||||||||| || |||| ||||||| ||| ||||| |||||||||| :||||||
Db    64  VDELIRVAKHSGADAVIPGYGFGFLSENPDLAAKCAEAGITVGPSAEVLQLTGNKARAIAA 123
Qy   134  AKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKLATE 193
          |: ||||| | ||| ||     :::|||::| |||||||||||||||| : : |
Db   124  ARAAGLPVLSSSSEPSSSSVDELMAAAADMEFPLFVKAVSGGGRGMRRVTDRESLAEAIEA 183
```

```
Db   194  ASREAEAAFGDGAVYVERAVINPQHIEVQLLGDHIGEVVHLYERDCSLQRRHQKVVEIAP 253
          ||||||| ||||||| ||||||||| ||| ||| |||| |||  |   ||||||||||||
Qy   184  ASREAESAFGDASVYLEQAVLNPRHIEVQILADGAGNVMHLFERDCSVQRRHQKVVELAP 243
Qy   254  AQHLDPELRDICADAVKFCRSIGYQGAGIVGEFLVDEKGNHVFIEMNPRIQVEHTVTEEV 313
          | :| ||| ||||| ||  : || ||||| |||| ||| |||||||||||| ||  ||:
Db   244  APNLSDELRQQICADAVAFARQIGYSCAGIVEFLLDERGHVFIECNPRIQVEHTVTEEI 303
Qy   314  TEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAYRS 373
          | ||| ::|| | ||| ||||::| :||| ||  | :|||||||| ||||||||||
Db   304  TDVDLVSSQLRIAAGETLADLGLSQDRLVVRGAAMQCRITTEVPANGFRPDTGRITAYRS 363
Qy   374  PGGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIG 433
          |||||||||| : ||| ||||||| |:||| ||| |  ||| || |||||
Db   364  PGGAGIRLDGGTNLGAEISAHFDSMLVKLTCRGRDFSAAASRARRALAEFIRGVSTNIP 423
Qy   434  FLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVNKPHGVRPKDV 493
          |: |: |   ||:|| :|   ||   ||  |   | | |||: ::||||| | |:|
Db   424  FLQAVIDDPDFRAGRVTTSFIDDRPHLLTSRSPADRGTRILNYLADITVNKPHGERPSTV 483
Qy   494  AAPIDKLPNIKDL--PLPRGSRDRLKQLGPAAFARDLREQDALAVTDTTPRDAHQSLLAT 551
          |:|||:| :|    ||| ||  |    | ||| :|| ::  | ||| |||||| ||||||
Db   484  -YPQDKLPPL-DLQAPPAGSKQRLVELGPEGFAGWLRESKAVGVTDTTPRDAHQSLLAT 541
Qy   552  RVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVNIQ 611
          || |  || | ||| |  :|:| ||||||| ||||:|  |:|| ||  || |  | |  |
Db   542  RVRTTGLLMVAPYVARSMPQLLSIECWGGATYDVALRFLKEDPWERLAALRESVPNICLQ 601
Qy   612  MLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETNTAV 671
          |||||||||||||| | :|||||||| :|||||||||||||||| || | : |
Db   602  MLLRGRNTVGYTPYPELVTSAFVEEAAATGIDIFRIFDALNNVESMRPAIDAVRETGSTI 661
Qy   672  AEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLVTAL 731
          |||| || ||  || |||||:|||||||| : :|||  | |||||||||||  ||||:||
Db   662  AEVAMCYTGDLSDPAENLYTLDYYLKLAEQIVEAGAHVLAIKDMAGLLRPAAHTLVSAL 721
Qy   732  RREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFAHT 791
          | | ||||||||||||||||||| ||| ||||||||||| |||||||||| ||||||
Db   722  RSRFDLPVHVHTHDTPGGQLATYLAAWSAGADAVDGASAPMAGTTSQPALSSIVAAAHT 781
Qy   792  RRDTGLSLEAVSDLEPYWEAVRGLYLFFESGTPGPTGRVYRHEIPGGQLSNLRAQATALG 851
          | | || |  ||||||| ||||||| : ||||||||||| |||||||||||||| |||
Db   782  QYDTGLDLRAVCDLEPYWEAVRKVYAPFESGLPGPTGRVYTHEIPGGQLSNLRQQAIALG 841
Qy   852  LADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAADPQKYDIPD 911
          | ||| || | ||| ||| |||||||||| ||||||||||||||||  |||||||||||
Db   842  LGDRFEEIEANYAAADRVLGRLVKVTPSSKVVGDLALALVGAGITAEEFAEDPAKYDIPD 901
Qy   912  SVIAFLRGELGNPEPGGWPEPLRTRALEGRSEGKAPLIEVPEEQAHLDADDSKERRNSLN 971
          |||  ||||| ||| ||||||||||| ||  || |  ||   |      :||:|||||
Db   902  SVIGFLRGELGDPPGGWPEPLRTRKALQGRGPAR-PVEKLTADDEALL-AQPGPKRQAALN 959
Qy   972  RLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISEP 1031
          |||| | || |||:|  || ||:: |||||| || | ||  :|| : |  :|
Db   960  RLLFPGPTAEFEAHRETYGDTSSLSANQFFYGLRYGEEHRVQL-ERGVELLIGLEAISEA 1018
Qy  1032  DDKGMRNVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAGVVTVAEGD 1091
          ::| || ::|| | ||  |  |||||||:||  | |:||:|:|||||||||||||||||
Db  1019  DERGMRTVMCIINGQLRPVLVRDRSIASEVPAAEKADRNNADHIAAPFAGVVTVGVAEGD 1078
Qy  1092  EVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLVVVS 1140
          |:| |||||||||||||||| :||:|| |||| | | |||||||||||
Db  1079  SVDAGQTIATIEAMKMEAAITAPKAGTVARVAVATAQVEGGDLVVVS 1127
```

```
RESULT 7
P95127
ID   P95127
AC   P95127;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
```

```
DE   Hypothetical 120.4 kDa protein (Pyruvate carboxylase).
GN   PCA OR RV2967C OR MTCY349.20 OR MT3045.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37Rv;
RX   MEDLINE=99295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z83018; CAB05410.1; -.
DR   EMBL; AE007125; AAK47371.1; -.
DR   HSSP; P24182; 1BNC.
DR   TIGR; MT3045; -.
DR   TubercuList; Rv2967c; -.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   InterPro; IPR000634; S/T_dehydrtse.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; Pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
DR   PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   1127 AA;  120422 MW;  84B0A4CC1A23CD90 CRC64;

Query Match             64.1%;  Score  3712.5;  DB 16;  Length 1127;
Best Local Similarity   64.4%;  Pred. No. 6.3e-191;
Matches  730;  Conservative  153;  Mismatches  239;  Indels  11;  Gaps  6;

Qy     12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 71
           |:|||||||||:|||||| |  | ||||| |  |  ||| | :|| |||
Db      2 FSKVLVANRGEIAIRAFRAAYELGVGTVAVYPYEDRNSQHRLKADESYQITGIGHPVHAY 61

Qy     72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTIGDKSRAV 131
           |::|: ||| :|| ||||||||||||||   |   |  ||| ||:|||:|:|
Db     62 LSVDEIVATARRAGADAIYPGYGFLSENPDLAAACAAGISFVGPSAEVLELAGNKSRAI 121

Qy    132 TAAKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKLA 191
           |:|:||| | | ::: | | :|:||||||||||||| |

Db    122 AAAREAG:PVLMSSAPSASVDELLSVAAGMPFPLFVKAVAGGGRGMRVGDIAALPEAI 181
           |||: | |||:||:||:||::|::|| :|| :|:||||||||:|||||:| |  |
Qy    192 TEASREAEAAFGDGAVYYERAVINPQHIEVQILGDHTGEVHLYERDCSLQRRHQKVVEI 251
           |||||||:||||| |:  |||||||:||||||| |: :||:||||||:||||||||:||
Db    182 EAASREAESAFGDPTVLEQAVINPRHIEVQILADNLGDVIHLYERDCSVQRRHQKVIEL 241

Qy    252 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE 311
           ||| |||:|||  :| ||| | | ||| |||||||:|::  ||||||||||:|||||||
Db    242 APAPHLDAELRYKMCVDAVAFARHIGYSCAGTVEFLLDDERGEVVFIEMNPRVQVEHTVTE 301

Qy    312 EVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAY 371
           |:|:||||::| ||||| ||::|||:|| |  ||||||||||||||:|||||||| |:|
Db    302 EITDVDLVASQLRIAAGETLEQLGLRQDIAPGAALQCRITTEDPANGFRPDTGRISAL 361

Qy    372 RSPGGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATN 431
           |: ||||||||:: :|| | |:|||||| ||||| :  || ||:|| |||  |||||:|||
Db    362 RTAGGAGVRLDGSTNLGAEISPYFDSMLVKLTCRGRDLPTAVSRARRAIAEFRIRGVSTN 421

Qy    432 IGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADPDEQGRIILDYLADVTNKPHGVRPK 491
           | ||:|:| :: ||:  || | ||| |:|| | | :|  | |: |:||||||:  || ||
Db    422 IPFLQAVLDDPDFRAGRVTTSFIDERPQLLTARASADRGTKILNFLADVTVNNPYGSRPS 481

Qy    492 DVAAPIDKLPNIKDLPL------PRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDAHQS 547
           :   |||||    || |      ||:|| ||::||| :|||:|||  | |||||||||||||
Db    482 TI-YPDDKLP---DLDLRAAPPAGSKQRLVKLGPEGFARWLRSAAVVTDTTFRDAHQS 537

Qy    548 LLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPN 607
           ||||||| | | | ||| || || ||||| ||||||||| ||| ||||:|| ||| |||
Db    538 LLATRVRTSGLSRVAPYLARTMPQLLSVECWGGATYDVALRFLKEDPWERLATLRAAMPN 597

Qy    608 VNIQMLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVLET 667
           : :|||:|||||||||||: | :|||:||::||||||||||||||:: :||||||| |
Db    598 ICLQMLLRGRNTVGYTPYPEIVTSAFVQEATATGIDIFRIFDALNNIESMRPAIDAVRET 657

Qy    668 NTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKL 727
           ::|:|||| | |||:|| |::|||||::|:|| :||  ||| |||||||||||| |:|
Db    658 GSAIAEVAMCYTGDLTDPGEQLYTLDYYIKLAEQIVDAGAHVLAIKDMAGLLRPPAAQRL 717

Qy    728 VTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADVDGASAPLSGTTSQPSLSAIVAA 787
           |:|||  |||||| |||||:|||||:|:||  | |||||| | ||  |||| ||:|||||||
Db    718 VSALRSRFDLPVHLHTHDTPGGQLASYYAAWHAGADVDGAAAPLAGTTSQPALSSIVAA 777

Qy    788 FAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSNLRAQA 847
           :|||  ||||| :|  |||||||:||:|  ||||| |||||||| ||||||||||||:||
Db    778 AAHTEYDTGLSLSAVCALEPYWEALRKVYAPFESGLPGPTGRVYHHEIPGGQLSNLRQQA 837

Qy    848 TALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHIVGAGYDPADFAAADPQKY 907
           ||||  ||||:| : |:| :: | | |||||:|||||||||:||||| : |:|:: ||::
Db    838 IALGLGDRFEEITEEAYAGADRVLGRLVKVTPTSKVVGDLALALVGAGVSADEFASDPARF 897

Qy    908 DIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEGAHLDADDSKERR 967
           ||| ||: |||||||  |||||||| | | || ::  | ||: |: :|   :| |: ||
Db    898 GIPESVLGFLRGELGLGDPPGGWPEPLRITAALAGRGAAR-PTAQLAADDEIALSSVGAK-RQ 955

Qy    968 NSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDA 1027
           ::||||||| ||:||::|| | | ||:| :: :|||| ||| | ||     |:: | |
Db    956 ATLNRLLFPSPTKEFNEHREAYGDTSQLSANQFFYGLRQGEEHRVKL-ERGVELLIGLEA 1014

Qy   1028 ISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAGVVTVV 1087
           |||||::||  |:  ::|| :||: |||||:: | |||:||  | || |||||||||| |
Db   1015 ISEPDERGMRTVMCILNGQLRPVLVRDRSIASAVPAAEKADRGNPGHIAAPFAGVVTVGV 1074

Qy   1088 AEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS 1140
           : ||:| || : | |||||||| |||| :| :  ||| :| ||| ||:|||||
Db   1075 CVGERVGAGQTIATIEAMKMEAPITAPVAGTVERVAVSDTAQVEGGDLVVVVS 1127

RESULT 8
Q50450
ID   Q50450       PRELIMINARY;      PRT;    1124 AA.
AC   Q50450;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLre. 01, Last sequence update)
```

```
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyc.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Smith D.R.;
RL   Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Robison K.;
RL   Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- COFACTOR: BIOTIN (BY SIMILARITY).
DR   EMBL; U00024; AAA50948.1; -.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   InterPro; IPR000534; S/T_dehydrtse.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   TIGRFAMs; TIGR01235; Pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
DR   PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW   Biotin.
SQ   SEQUENCE   1124 AA;  120672 MW;  07E1692E12203491 CRC64;

Query Match             60.2%;  Score 3483;  DB 2;  Length 1124;
Best Local Similarity   62.2%;  Pred. No. 1.3e-178;
Matches  698;  Conservative  151;  Mismatches  251;  Indels   22;  Gaps    9;

Qy    12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 71
            |::||||||||||:|||||| |||:|| |:| |||:  :::: ||| ||:|||   | |
Db     2 FSKVLVANRGEIAIRAFRAAYELGVGTVAVYPYEDRNSQHRLKADESYQIGDIGHPVHAY 61

Qy    72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 131
         | :|||: ||:: ||||||||||||||| |||.|||.||:|:||: ||| | |:|||||
Db    62 LSVDEIVATARRAGADAIYPGYGFLSENPDLAAACAAAGISFVGPSAEVLEAGNKSRAI 121

Qy   132 TAAKKAGLPVLAESTPSKNIDEIVKSARGQTYPIFVKAVAGGGGRGMRFVASPDELRKLA 191
           |:::|||||:: |:|: ::|::|| |: |||:||||||||||||||||| |:  |  |
Db   122 AAAREAGLPVLMSSAPSASVDELLSVAAGMPFLFVKAVAGGGGRGMRVGDIAALPEAI 181

Qy   192 TEASREAEAAFGDGAVVVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVEI 251
          |||||||::||||:||  |||||||:|||||||:|  |||||||||||:||||||||:|
Db   182 EAASREAESAFGDPTVYLEQAVINPRHIEVQILADNLGDVIHLYERDCSVQRRHQKVIEL 241

Qy   252 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE 311
         || |:||.||  |:| |||:|| :|||: |||||||:|::|: ||||||||:|||||||||
Db   242 APAPHLDAELRYKMCVDAVAFARHIGYSCAGTVEFLLDERGEYVFIEMNPRVQVEHTVTE 301

Qy   312 EVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAY 371
         |:|:||||: |:|:::||.||:|||  : | ||||||||||||||| |||||  | :||
Db   302 EITDVDLVASQLRIIAAGETLEQLGLGLRQEDIAPHGAALQCRITTEDPANGFRP-TRAGSAR 360

Qy   372 RSPGAGVRLDGAAQLGGEITAHFDSMLVKM----TCRGSDFETAVARAQRALAEFTVS 426
          .:   :    |::  ::    :  :::::    |  : | | |||:||:|||||| :
Db   361 CDPPAVPVSAWTAAP-----TWRRNQPVLRLHAGQADLSGRDLPTAVSRARRAIAEFRIR 415

Qy   427 GVATNIGFLRALLREDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVNKPH 486
         |::|||||| ||::| :::  :| ||:|||| ||||: ||||:  || :|||||||:||
Db   416 GVSTNIPFLQAVLDDDFDRAGRVTTSFIDERPQLLTARASADRGIKILNFLADVTVNNPY 475

Qy   487 GVRPKDVAAPIDKLPNIKDLPL----PRGSRDRLKLQLGPAAFARDLREQDALAVTDTFR 542
         | ||  : ||::::| ::|||     | ||: || ||||  ||  |||:::| ||||||
Db   476 GSRPSTI-YPDDKLP---DLDLRAAPPAGSKQRLVKLGPEGFARWLRESAAVGVTDTFR 531

Qy   543 DAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELR 602
         |||||||||||| :|| | :||:|:| ||||||| ||||:|||| ||| ||| || | ||
Db   532 DAHQSLLATRVRTSGLSRVAPYLARTMPQLLSVECWGGAITYDVALRFLKEDPWERLATLR 591

Qy   603 EAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAID 662
         .|||| ::||||||||||||||||::|  :||:||::.|:|||||||||| ::|:||||
Db   592 AAMPNICLQMLLRGRNTVGYTPYPEIVTSAFVQEATATGIDIFRIFDALNNIESMRPAID 651

Qy   663 AVLEINTAVAEVAMAYSGDLSDPNEKLYTLDYLKMAEEIVKSGAHILAIKDMAGLLRPA 722
         || | : :|||||| |:||| || ::|||||| |:|| || .||:|||||||||||||
Db   652 AVRETGSAIAEVAMCYTGDLTDPGEQLYTLDYYLKLAEQIVDAGAHVLAIKDMAGLLRPP 711

Qy   723 AVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLS 782
          :| :|::||.|||||||:|||||| ||||:| || :||||||||:: ||:|||||| ||
Db   712 AAQRLVSALRSRFDLPVHLHTHDTPGGQLASYVAAWHAGADAVDGAAAPLAGTTSQPALS 771

Qy   783 AIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLLPFESGTPGPTGRVYRHEIPGGQLSN 842
          ||||||:|| |||||| ||  |||||||| |:| ||||||:|||||||| |||||||||
Db   772 SIVAAAAHTEYDTGLSLSAVCALEPYWEALRKVYAPFESGLPGPTGRVYHHEIPGGQLSN 831

Qy   843 LRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVPADFAA 902
         || ||||||.||||:||: |||.:||||||||:||:|||||||||| ||||||.||: |
Db   832 LRQQATALGLGDRFEEIEEAYAGADRVLGRLVKVTPTSKVVGDLALALVGAGVSADEFAS 891

Qy   903 DPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDADD 962
         ||::: |||||: |||||||:|||||||||||| ||||  :  |  |  ::| |  ||
Db   892 DPARFGIPESVLGFLRGELGDPPGGWPEPLRTAALAGRGAAR-PTAQLAADDEIALSSVG 950

Qy   963 SKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLL 1022
          :: : :||||||| ||:||:|||.:.|:||:|. :::||||: :|| | :|:  | ||
Db   951 AK-RQATLNRLLFPSPTKEFNEHREAYGDTSQLSANQFFYGLRQGEEHRVKL-ERGVELL 1008

Qy  1023 VRLDAISEPDDKGMRNVANVNGQIRPMRVRDRSVSVTATAEKADSSNKGHVAAPFAGV 1082
         : ::|||||||:||| || :|||:|| ||||||: :| |||||| .|.|||||||||
Db  1009 IGLEAISEPDDERGMRTVMCILNGQLRPVLVRDRSIASAVPAAEKADRGNPGHIAAPFAGV 1068

Qy  1083 VTVTVAEGDEVKAGDAVAIEAMKMEATTIASVDGKIDRVVV 1124
         ||| | |::| ||| |:|:|||||||| || ||| |:||:|
Db  1069 VTVGVCVGERVGAGQVIATIEAMKMEAPITAPVAGVVERVAV 1110
```

RESULT 9
Q8JHF6

```
ID   Q8JHF6       PRELIMINARY;        PRT;  1178 AA.
AC   Q8JHF6;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase (EC 6.4.1.1).
GN   PYC.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22145742; PubMed=12150961;
RA   Jitrapakdee S., Nezic M.G., Cassady A.I., Khew-Goodall Y.,
RA   Wallace J.C.;
RT   "Molecular cloning and domain structure of chicken pyruvate
RT   carboxylase.";
RL   Biochem. Biophys. Res. Commun. 295:387-393(2002).
DR   EMBL; AF509529; AAM92771.1; -.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
```

```
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
KW   Ligase.
SQ   SEQUENCE   1178 AA;   127253 MW;   F0722FEAB8BF39A5 CRC64;

Query Match             44.1%;  Score 2552;  DB 13;  Length 1178;
Best Local Similarity   48.1%;  Pred. No. 1.5e-128;
Matches  556;  Conservative 165;  Mismatches 410;  Indels  26;  Gaps  14;

Qy     2 STHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRI 61
            ::!||::||!||||||| |::: :||||| ||  |  ||::|
Db    28 SVRSASCQP-IRKYLVANRGEIAIRVFRACTELGLRTVAVYSEQDTGQMHRQKADEAYLV 86

Qy    62 GTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVL 121
           :/|||||||| |||| ::  :|| |: ||||||| |:  : |:|||||:|||| : |
Db    87 GRGLJPVQAYLHVPDIIVARENAVDAIHPGYGFLSERADFAQACVDAGVRFVGPPPEVV 146

Qy   122 DLTGDKSRAVTAAKKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180
           || |: |||| |  || :|   ||| :||:|| |||   ||| ||||||| |||||||
Db   147 RKMGDKVEARSIAIAAGVPVVPGTSAPVATLGEAQDFAARVGFPIIFKAAHGGGGRGMRA 206

Qy   181 VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240
           | |: :|  |:  || | |||||| || |  ||||  ||||:|| |||  ||||||||
Db   207 VRGPQELESFSRASSEALAAFGDGALFVEKLMERPRHIEGQILGDKHGNVVHLYERDCS 266

Qy   241 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGTYQGAGTVEFLVDEKGNHVFIEMN 300
           :/||||||||||||||:||| |::  : :||  : :|:|  || |||||:||::|||||
Db   267 IQRRHQKVVEIAPAARLDPQLRAQLASDAVRIAQQVGYENAGTVEFLVDRDGKHYFIEVN 326

Qy   301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
           |:|||||||| |:|| :|| |:|| :  : :|:   ||||:|||||||| |||:||||:
Db   327 SRLQVEHTVTEEITGVDLVQAQLVAAGRSLSEIGLQQDSVRVNGCAIQCRVTTEDPARG 386

Qy   361 FRPDTGTTTAYRSPGGAGVRLDGAAQL-GGEITAHFDSMLVKMTCRGSDFETAVARAQRA 419
           | ||:||||   || | | |:|||| : | ||||||||:||||  |||||| ||||||
Db   387 FQPDTGRIEVFRSGEGMGIRLDGASAFQGALISPHYDSLLVKVIAHGPDQPSAAAKMSRA 446

Qy   420 LAEFTVSGVATNIGFLRAILREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLAD 479
           :/||||| | ||||||||||||:|||||:||| ||:|| ||||   |::|||||||:||
Db   447 LGEFRIRGVKTNIPFLQNVLAHPQFLGGAVDTQFIDENPELFHLRPSQNRAQKLLHYLGH 506

Qy   480 VTVNKPHGVRP-KDVAAPIDKLPNIKDLPLPRGS-----RDRLKQLGPAAFARDLREQDA 533
           | |:|| | || |::||:  :|| ||: :|: |       ||||||:|| |||||:|||
Db   507 VMVNGPSTPLPVKAKAAVVEPVPP---PVPMGSPPEGLRAVLQREGPAGFARALRGHRG 562

Qy   534 LAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFED 593
           |/|||||||||||||||||||  | |:|||  |||:||||||  |||||||||||:| ||
Db   563 LLLXDTTFRDAHQSLLATRVRTRDLARIAPFVAHSLSPLCSMETWGGATFDVAMRFLHEC 622

Qy   594 PWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALND 653
           |/| || | ||| |:| |||||| | |  || :||| ||||:| ||||:||| || | |
Db   623 PWERLRELRRLVPNIPFQMLLRGANAVGYTNYPDNVIYRFCEVAAANGMDIFRIFDALNY 682

Qy   654 VSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIK 713
           |:|||||| |||||||||  ||| |||||:|  :| |||||||| :|:|| : |||||||
Db   683 LPNLLGVEAVGRAG-AVVEAALSYTGDVADPTRTKYSLDYYLGLAKELVAAGTHILCIK 741

Qy   714 DMAGLLRPAAVTKLVTALRREF-DLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPL 772
           |/|||||||| ||:||| ||   :||||||||||| || ||  || ||:||||||:| :
Db   742 DMAGLLTPAAARLLVSSLRDRFPDVPIHVHTHDTAGAIATLLAANADADVVDVAVDAM 801

Qy   773 SGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGR--V 830
           |/|||||||| || || |||||||||| || ||:||| |||||| |::| |||| ||   |
Db   773 SGTTSQPSMGALVACARGTPLDTGIALERVFEYSEYWEGARGLYAAFDCTATMKSGNADV 861

Qy   831 YRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHL 890
           |:||||||||| | |:|  | | :||||||:||  |:|:| |   ||||||||||| :|:
Db   862 YENEIPGGQYTNLHFQAHAMGLGHKFKEVKKAYAEANKLLGDLIKVTPSSKVVGDLAQFM 921

Qy   891 VGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEG--RSEGKAPLT 948
           ||||| || |:|||:|||| ||:  |:  || | || |||| ||| : : ||:  |||
Db   922 VQNGLSREEAEARADELSFPLSVVEFLQGYIGTPPGGFPEPFRSKVLKDLPRVEGRPGAS 981

Qy   949 EVPEEEQA---HLDADDS--KERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYG 1003
           |/||::|   | |  |::   ::||:||| |  |:| :| | |:|:| : ::|||:: |
Db   982 LPPLDFEALSQELGARDGTPPSPEDLLSAALYPKVYAEFRDFTSTFGPVSCLGTRLFLEG 1041

Qy  1004 LVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTAT 1063
           :/|  :| | |:|| | :  |  || |||||::| |:| |     ||: ||| :|
Db   1042 PTIAEEFEVELERGKT-LHIKALGLDLNAAGQREAFFELNGQLRSIIVRDTQALKEMHV 1100

Qy  1064 AEKADSSNKGHVAAPPAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRV 1122
           |:|||: |||| | | |  ||:||||||||||||| || :||||||  | |:|||||||
Db  1101 HPKADRSAKGQVGAPMPGEVVEVRVKEGEAVEKGAPLCVLSAMKMETVVTAPRGGTVSRL 1160

Qy  1123 VVPAATKVEGGDLIVVV 1139
           | | | |||||| ||||
Db  1161 HVRPGMSLEGDDLIAEI 1177

RESULT 10
Q9K9M0
ID   Q9K9M0                 PRELIMINARY;       PRT;   1150 AA.
AC   Q9K9M0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase.
GN   PYCA OR BH2625.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
DR   EMBL; AP001516; BAB06344.1; -.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   1150 AA;   128858 MW;   ED6788BE8A9F3BD4 CRC64;

Query Match             43.6%;  Score 2525.5;  DB 16;  Length 1150;
Best Local Similarity   46.8%;  Pred. No. 3.9e-127;
Matches  539;  Conservative 185;  Mismatches 400;  Indels  27;  Gaps  13;

Qy     9 LPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPV 68
```

```
Db        1077  SNPNHIGASMPGTVKALVEKGDKVKQGDHLMITEAMKMETVQAPFDGEVVALHVKDGD 1136

Qy        1129  KVEGGDLIVVV 1139
                :::|||:|::|
Db        1137  AIQTGDLLIEV 1147
```

RESULT 11
Q9KWU4

```
ID   Q9KWU4;
AC   Q9KWU4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase (EC 6.4.1.1).
GN   PYCA.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida H., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z99111; CAB13359.1; -.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
```

```
Qy           4  LKNIKKVLVANRGEIAIRIFRACTELHIRTVAISKEDIGAYHRYKADEAYLVGEGKKPI 63
                 |::||||||||||:||| ||:|||||:|  ||| | :|| | ||||| :|| ::  |:
Db          69  KAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLIGDKS 128
                : ||||| :||  ||: ||:|||||||||::|  :| ||| |||:|||| ||||||:|||
Db          64  EAYLDIEGIIEIAKRHGVDAIHPGYGFLSENIEFAKRCHEGIIFIGPELEHLVMFGDKV 123

Qy         129  RAVIAAKKAGLPVLAEST-PSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRFVASPDEL 187
                 |::|  | ||| | || |||:||: | || |::| |:||||  |||||:||| ||| :
Db         124  QAREQAIKANLPVIPGSDGVSSLEDVKAFADKHGYPFIIKAALGGGGRGMRIVRSENDV 183

Qy         188  RKLATEASRAEAAFGDGAYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQK 247
                 |:||:|::: || ||:|||:  |:|||:|||| :|||||| :|||||||: ||||:
Db         184  QESYERAKSEAKAAFGNDEVYVEKFIENPKHIEVQILADKHGNTLHLYERDCSVQRRHQK 243

Qy         248  VVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEH 307
                 ||::| :||| |:|:|  ::   : | |:||:::||||  :|:|| |||||||||||||
Db         244  VVEVAPSVSLSEDVRERICAAVQLAENVNYVNAGTVEFLVDREGNFYFIEVNPRIQVEH 303

Qy         308  TVTEEVTEVDLVKAQMRLAAGATL-KELGL-TQDKITHGAALQCRITTEDPNNGFRPD 364
                 |:| : || ||:    |  |||| |:||| |||:|||||||  || ||||||:||||
Db         304  TITEMVTGIDIVQSQLFIADGEHLHGDRLGIPKQEIVCHGYAIQSRVTTEDPSNGFLPD 363

Qy         365  TGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEF 423
                 || ||||| ||| | |:|  |:||| |||:|| || ::: :|:|| :| |||: | |||
Db         364  TGRINAYRSGGGFGVRLDAGNGFQGAVITPYYDSLLVKVSTWALTFEGAAAKMLRNLREF 423

Qy         424  TVSGVATNIGFLRALLREDDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVN 483
                 || ||:|| |:| |||  || :|| ||| : ||  : ::|| | |: ||||::||:||:
Db         424  RIRGIKTNIAFLENVVQHRQFLSGEYNTSFIDQTPELFVFPKRKDRGTKMLSFIGETIVN 483

Qy         484  KPHGVRPKDVAAPIDK--LPNIK-DLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 540
                 |  | ||| :  | |  || ||  |:|  || | ||:||  : |||:|| |:|  |||||
Db         484  GYPGLE-KTKKPVFDKPPVPKLKLSEPIPDGTKQILDQHGPEGLAKWVKEQKHVLLTDTT 542

Qy         541  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600
                 |||||||||||| |   |: :| :| :|||| || |||||||||||||||||:|   ||:
Db         543  FRDAHQSLLATRVRTHDLKQIAEPTARLLPNLFSMEMWGGATFDVAMRFLHEDPWERLLI 602

Qy         601  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 660
                 ||:| |:|:| ||||||||| ||||||| ||| |||:| ||||||||| |||::: ||
Db         603  LRKKAPNVLFQMLLRASNAVGYKNYPDNLIREFVDKSANAGIDVFRIFDSLNWVEGMKLA 662

Qy         661  IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR 720
                 ||: |||: :|: |  :| ||   || || |||| ||:|| :||||||||||||||||||
Db         663  IEAVGEAN-KIAEATICYTGDILDSSRPKYDLAYYKKLAKELEAGAHILGIKDMAGLLK 721

Qy         721  PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQGADAVDGASAPLSGTTSQPS 780
                 |||:||| || |||||||||||||||||||:||||  | |||:  |||| :|:|:|||:
Db         722  PEAAYQLVAELKDTVTIPVHLHTHDTSGNGIFTYARAIEAGVDIVDVAVSSMAGLTSQPS 781

Qy         781  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLFPESGTPGPTGRVYRHEIPGGQL 840
                 |||:|| || ||:|:| |  : :| | : || ||:|  |:|| || |  ::| | ||| :
Db         782  ANSLYYALADSERQPNVNITALEQLAEFWEETRKFYAGFESGMNAPHTEVYEHEMPGGQY 841

Qy         841  SNLRAQATALGLADRFELIEDNYAAVNEMLGRP TKVTPSSKVVGDLALHLVGAGVDPADF 900
                 |||:|||| ||| || |: : ::  |:||:::  ||:|||:||||||: :: : :|:| |
Db         842  SNLQQQAKAVGLGHRWNEVKKMYRTVNDMFGDVVKVTPSSKVVGDMALYMVQNDLTEEEV 901

Qy         901  AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA 960
                 | :  ||| ||||| || || |||||| |  :  ||| |::| |||||| |:||: : |:
Db         902  YENGHKLDFPDSVVEFFEGGLGQPYQGFPKKLQEIILKGRK---PITNRPGENMEPIQF 957

Qy         961  DDSKER-RNSLNR------------LLFPKPTEEFLEHRRFGNTSALDDREFFYGLVEGRE 1009
                 : :|:  |||::             | | |:|| ||:|||: |:: ||::::|||:| | |
Db         958  EAIKEELYNKLDRQVTSHDILSYALYPKVFMEFERFRQITFGDVSVLDITPTFFYGLRPGEE 1017

Qy        1010  TLIRLPDVRTPLLVRLDAISEPDKGMRNVANVNGQIRPMRVRDRSVESVTAEKADS 1069
                 |:: | ::||: |  |:  : :| ::      |:| | :   ::|:: :|  :|  |
Db        1018  IEVEIEQGKT-LIVKFISLSKPQDDGNRIVYFELNGQPREVLIKDQSVKTSIISRPKADK 1076

Qy        1070  SNKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAAT 1128
                 |  | |  | || : |  :|||:||| || | |:| ||||| | |||  || :|| |
```

```
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMS; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Ligase; Complete proteome.
SQ   SEQUENCE   1148 AA;  127936 MW;  98D14DD56F361620 CRC64;

Query Match             43.6%;  Score 2524;  DB 16;  Length 1148;
Best Local Similarity   46.0%;  Pred. No. 4.7e-127;
Matches  533;  Conservative  186;  Mismatches  392;  Indels  48;  Gaps  13;

Qy     11 AFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKA 70
           ::|:|||||||||||:|::|||:||::||::|||||:||||::|||:|||:|||
Db      5 SIQKVLVANRGEIAIRIFRACTELNIRIVAVYSKEDSSSYHRYKADEAYLVGEGKKPIDA 64

Qy     71 YLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRA 130
           ||||:||:||:||||||||:|||||||||:|:|||||:|||:||||::||||::|||:|:
Db     65 YLDIEGIIDIAKRNKVDAIHPGYGFLSENIHFARRCEEEGIVFIGPKSEHLDMFGDKVKA 124

Qy    131 VTRAKKAGLPVLAEST-PSKNIDEIVKSAEGQYPIFVKAVAGGGGRGMRFVASPDELRK 189
           :|||:|||:||::|:::||:||:||::||||:||||:|||:|||||||||:|:::|:|:
Db    125 REQAEKAGIPVIPGSDGPAETLEAVEQFGQANGYPIIIKASLGGGGRGMRIVRSESEVVKE 184

Qy    190 LATEASREAEAAFGDGAVYVERAVINPOHIEVQLLGDHTGEVVHLYERDCSLQRRHQKVV 249
           :||:||:||:|||:|:||:||::|:||:||||||:||::|:||||:|||||:|||||||:
Db    185 AYERAKSEAKAAFGNDEVVVEKLLENPKHIEVQVIGDKQGNVVHLFERDCSVQRRHQKVI 244

Qy    250 EIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTV 309
           |:||:::|:|||||:||:||||::|:::|:::|||||||:|::|:|:||:|||:||||||:
Db    245 EVAPSVSLSPELRDQICEAAVALAKNVNYINAGTVEFLV-ANNEFYFIEVNPRVQVEHTI 303

Qy    310 TEEVTEVDLVKAQMRLAAGATL-KELGLTQDK-IKTHGAALQCRITTEDPNNNGFRPDTG 366
           ||::|:||:|::|:::|:|::|::|::::::::|:|::::|:|:|:||||||::|:::|||
Db    304 TEMITGVDIVQTQILVAQGHSLHSKKVNIPEQKDIFTIGYAIQSRVTTEDPQNDFMPDTG 363

Qy    367 TITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTV 425
           :|:||||:|::|:|||:|::::|::::||:||:|||::|:::|::::||::|::::|:||:
Db    364 KIMAYRSGGSGFGVRLDTGNSFQGAVITPYYDSLLVKLSTWALTFEQAAAKMVRNLQEFRI 423

Qy    426 SGVATNIGFLRALLREEDFTSKRRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVN- 483
           :|::|||:||::::::|:|::::::|:||:|::::::::::::::::||:::|::|||||:
Db    424 RGIKTNIPFLENVAKHEKFLTGQYDTSFIDTTPELFNFPKQXDRGTKMLTYIGNVTVNGF 483

Qy    484 ------KPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAV 536
                 ::|:::|:::::::::|::|::::::::::|::::::::::|::::::|::::::
Db    484 PGIGKKEKPAFDKPLGVKVDVDQQP------ARGTKQILDEKGAEGLANWVKEQKSVLL 536

Qy    537 TDTTFRDAHQSLLATRVRSFALKPAAEAVALTPELLSVEAWGGATYDVAMRFLFEDPWD 596
           |||||||||||||||||:||::||::||:||:|||:|:|:|||||:||||:|||:|||:
Db    537 TDTTFRDAHQSLLATRIRSHDLKKIANPTAALWPELFSMEMWGGATFDVAYRFLKEDPWK 596

Qy    597 RLDELRFAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIPFDALNDVSQ 656
           ||::||::::|:::|||||::|::|||:|||:|::::|||:|::||:|:||||||:||:::
Db    597 RLEDLRKEVPNTLFQMLLRSSNAVGYTNYPDNVIKEFVKQSAQSGIDVFRIPFDSLNWVKG 656

Qy    657 MRPAIDAVLETNTAVAVEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMA 716
           |:::||||::|:::|:::::|:||:|::::::|:|:|:||::|||:::::||||||:||||
Db    657 MTLAIDAVRDTG-KVARAAICYTGDILDKNRTKYDLAYYTSMAKELEAAGAHILGIKDMA 715

Qy    717 GLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTT 776
           |||:|::|::||:||::|:|:|||:|||||:|::::||::|:::||||::|||:::::|:|
Db    716 GLLKPQAAYELVSALKETIDIPVHLHTHDTSGNGIYMTAKAVEAGVDIIDVAVSSMAGLT 775

Qy    777 SQPSLSAIVAAFAHTRRDTGLSEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIP 836
           ||||:|::::::::::|:::::::|::::::|||:|::||:::|||::::::::|:||:|
Db    776 SQPSASGFVHAMEGNDRRPEMNVQGVELLSQYWESVRKYYSEFESGMKSPHTEIYEHEMP 835

Qy    837 GGQLSNLRAQATAIGLGADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVD 896
           |||:|||::::::::||:|::::::::::||::::::::::|||||||||::||:::::::
Db    836 GGQYSNLQQOAKGVGLGDRWNEVKEMYRRVNDMFGDIVKVTPSSKVVGDMALIYMVQNNLT 895

Qy    897 PADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEE--- 953
             ::::::::::||||||:::|:|::|:|:||:||:|:::::::|::||:|::|::|
Db    896 EKDVYEKGESLDFPDSVVELFKGNIGQPHGGFPEKLQKLIIKGQE----PITVRPGELLE 951

Qy    954 --------EQAHLDADDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFF 1001
                   |::::|:|:::::::||:|:||::::|:|::::::::||||:||||:|::
Db    952 PVSFRAIKQEFKEQHNLEISD---QDAVAYALYPKVFTDYVKTTESYGDISVLDTPTFF 1007

Qy   1002 YGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVT 1061
           ||::||::::|:|||::|:|::::|::|||::::|:||:::||:|||:::|:::::::|:
Db   1008 YGMTLGEIEVEIERGKT-LIVKLLSIGEPQPDATRVVYFELNGQPREVVIKDESIKSSV 1066

Qy   1062 ATAEKADSSNKGHVAAPFAGVVTVTVAE-GDEVKAGDAVAIIEAMKMEATTASVDGKID 1120
           :::|||:::|::||:|:::|:|::::||:|::|:|||::::||||||||:|::::::||:
Db   1067 QERLKADRTNPSHIASMPGTVIKVLAEAGTKVNKGDHLMINEAMKMETTVQAPFSGTIK 1126

Qy   1121 RVVVPAATKVBGGDLIVVV 1139
           :::::|::::::|||::::
Db   1127 QVHVKNGEPIQTGDLLLEI 1145
```

RESULT 12
Q9DDT1

```
ID   Q9DDT1      PRELIMINARY;      PRT;    1180 AA.
AC   Q9DDT1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Pyruvate carboxylase.
GN   PC.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21100417; PubMed=11167010;
RA   Yoder J.A., Litman G.W.;
RT   "The zebrafish fth1, slc3a2, men1, pc, fgf3 and cycd1 genes define two
RT   regions of conserved synteny between linkage group 7 and human
RT   chromosome 11q13.";
RL   Gene 261:235-242(2000).
CC   -!- COFACTOR: BIOTIN (BY SIMILARITY).
DR   EMBL; AF295372; AAG37836.1; -.
DR   HSSP; P24182; 1BNC.
DR   ZFIN; ZDB-GENE-000831-1; pc.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPSase_L_D2.
DR   InterPro; IPR005481; CPSase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMS; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
KW   Biotin.
SQ   SEQUENCE   1180 AA;  129884 MW;  26118F2691400E24 CRC64;

Query Match             43.5%;  Score 2519.5;  DB 13;  Length 1180;
Best Local Similarity   47.2%;  Pred. No. 8.6e-127;
Matches  546;  Conservative  175;  Mismatches  414;  Indels  23;  Gaps  15;

Qy      2 STHTSS-TLP--AFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEA 58
           |:|:||:::|::::::|:||||||||:|:||||:||:::|:|||||:||:||||:||:||
Db     26 SAHLSSQTLEYHPIKKVMVANRGEIAIRVFRACTELGIRTVAVYSEQDTGQMHRQKADEA 85

Qy     59 VRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTP 118
```

Db   1162 HVNADSSLEGEDLILEIT 1179

RESULT 13
Q9HES8
ID   Q9HES8;   PRELIMINARY;   PRT;   1192 AA.
AC   Q9HES8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase (EC 6.4.1.1).
GN   PYC.
OS   Aspergillus niger.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX   NCBI_TaxID=5061;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N400;
RA   Panneman H., Ruijter G.J.G., Van den Broeck H.C., Visser J.;
RT   "Aspergillus niger pyruvate carboxylase.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- COFACTOR: BIOTIN (BY SIMILARITY).
DR   EMBL; AJ009972; CAC19838.1; -.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Biotin; Ligase.
SQ   SEQUENCE 1192 AA; 130866 MW; C7630C689BA23A64 CRC64;

Query Match   43.4%; Score 2509.5; DB 3; Length 1192;
Best Local Similarity 46.7%; Pred. No. 3e-126;
Matches 546; Conservative 175; Mismatches 404; Indels 45; Gaps 17;

Qy   2 STHT----SSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASE 57
            :|| : : |:||||||| |:||||| : : ||| :|:|: ::||:|:|||:|:||::|
Db   27 SVHTRLRANSAIMQFQKILVANRGEIPIRIFRTAHELSLQTVAVYSHEDHLSMHRQKADE 86

Qy   58 AVRIGTEG--SPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIG 115
            ||:| || |||| |||||:|||||| |||||||||||||||||||||||| ||:|||
Db   87 AYMIGKRGQYTPVGAYLAIDEIVKIALEHGVHLIHPGYGFLSENAEFARKVEQSGMVFVG 146

Qy   116 PTPEVLDTGDKSRAVTAAKKAGLPVLAEST-PSKNIDEIVKSAEGQTYPIFVKAVAGGG 174
            |||| || || ||:|||||||||| | |:| ||||||| |: || :||||:|||||||
Db   147 PTPQTIESLGDKVSARQLAIRCDVPVVPGTPGPVERYEEVKAFTDTYGFPIIIKAAFGGG 206

Qy   175 GRGMRFVASPDELRKLATEASREAEAAFGDGAVYVERAV-INPQHIEVQILGDHTGEVHL 234
            |||||| |:|:|||||:|:|:|: ::|: || :| |: | : | | : | | |
Db   207 GRGMRVVRDQAELRDSFERATSEARSAFGNGTVFVERFLDRPKHIEVQLLGDNHGNVVHL 266

Qy   235 YERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNH 294
            ||||||:||||||||||||||||||||||| |:|||||||||||||||||||||||||
Db   267 FERDCSVQRRHQKVVEIAPAKDLPADVRDRILADAVKLAKSVNYRNAGTAEFLVDQQNRY 326

Qy   295 VFIEMNPRIQVEHTVTEEVLEVDLVKAQMRLAAGATLKELGLTQDKIRTSRGFAIQCRITT 354
            :||| |||||||| |||:|| |||:|| |||: || |||||||||||:| :|| |||||||
Db   327 YFIEINPRIQVEHTITEEITGDIVAAQIQIAAGATLEQLGLTQDRISTRGFAIQCRITT 386

Qy   385 ...

Left column (Db/Qy alignment continued):

Db   86 YLIGRGLSPVAAYLHIPDIIKVAKENNVDAIHPGYGFLSERADFAQACAEAGVRFIGPSP 145
            ||:||:| |: :| ||||||||:|||| ||||||||||| | | |||||| |||:|
Qy   119 EVLDLTGDKSRAVTAAKKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRG 177
            |:| :: | || |||||| || :| ::|| |:|| | |||||||||||| ||||
Db   146 EVVRKMGDKVEARALAIKAGVPVVPGTDAPISCLQEAQEFAKTYDFPIIFKAAYGGGGRG 205

Qy   178 MRFVASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVHLYER 237
            || |: |:|||||:|:|:|: ::|: || :| |: | : | | : | | ||| |
Db   206 MRVVRNYEELEENYQRAYSEALAAFGNGALFVEKFIEKPRHIEVQILGDKYGNVIHLYER 265

Qy   238 DCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFIVDEKGNHVFI 297
            |||:|||||||||||||||||| | | | |:|||||||||||||||||||||||||| | :
Db   266 DCSIQRRHQKVVEIAPAAHLDSHLRDRLTLDSVNLAKQVGYENAGTVEFLVDKGKHYFI 325

Qy   298 EMNPRIQVEHTVTEEVLEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDP 357
            |:||||||||| |||:|| |:|:||||||||||||:|||||||:|:|||:|::| |||||
Db   326 EVNSRLQVEHTVTEITDVDLVHAQLRVCEGRSLPELGLEQDKIQINGCAIQCRVTTEDP 385

Qy   358 NNGFRPDTGTITAYRSPGGAGVRLDGAAQL-GGEITAHFDSMLVKMTCRGSDFETAVARA 416
            :||| | ||: |: ||| ||| ||| |:| || ::| |||:||:||| ||||::| :|
Db   386 SRGFQPDTGRIEVFRSGEGMGIRLDSASAFQGAIISPHYDSLLVKIASGKDLPTAATKM 445

Qy   417 QRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDY 476
            ||| ||| |||| ||||||:|| ::|:| :||:|| ||| |||:| |||:| |::||
Db   446 HRALTEFRVRGVKTNIPFLQNVLSNSQFLYATVDTQFIDENQNLFNLKPTQNRAQKLLHY 505

Qy   477 LADVTVNKPHGVRP-KDVAAPIDK-LPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDAL 534
            | |||| | || | ||| |:| |||:|||| | || ||: | | | : |:| :||
Db   506 LGHVMVNGPMTIPVKAKPSPVDPVIPSVSLGEPLGFREVLLREGPEGFARAIRQHQGL 565

Qy   535 AVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDP 594
            | |||||||||||||||||:|||| |: |||| |||:|:|||| |||||| ||| |||
Db   566 LLMDTTFRDAHQSLLATRVRTHDLKKIAPYVSHNFSNLFSLENWGGATFDAAMRFLSECP 625

Qy   595 WDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDV 654
            |:||:|||: || | | :| || ||| :|:||| |:| | ::||| || | ||:
Db   626 WKRLQELRALIPNVPFQMLLRGANAVGYTNYPDNAVFKFCEVAKENGMDIFRVFDSLNYI 685

Qy   655 SQMRPAIDAVLETNTAVEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKD 714
            | : | | :| | ||:|||| || ||:|| :| ||||| | ||||||| |:| |||
Db   686 PNMLLGMEAAGAAG--GVVEAAISYTGDVSDPMRQKYSLDYYRLKIADELVKAGTHILSIKD 744

Qy   715 MAGLLRPAAVTKLVTALRREF-DLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLS 773
            ||||| ||: ||:|:||||| :|:| |:|||||| :|| | ||: ||||||| |:|
Db   745 MAGLLKPQASRLLIEALRDRFPDIPIHVHTHDTAGAGVAAMLACAQAGADIVDVAVDSMA 804

Qy   774 GTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLLPFESGTPGPTGR--VY 831
            | ||| |||:|| :|||:| : | : :|: |:| | |:|||| |:|||| | | :|
Db   805 GMTSQPSMGAIVACTKGFKLDTGISLDKVFDYSEYWEVARGLYAPFDCTATMKSGNADVY 864

Qy   832 RHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHJV 891
            : |||||| || || | ||| | | |:::| | :| || |||||||:|| | | |:|
Db   865 ENEIPGGQYTNLHFQAHSMGLGNKFKEVKKAYTEANKLLGDLIKVTPSSKIVGDLAQFMV 924

Qy   892 GAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEG--RSEGKAPLTE 949
            || | | | |||::| || |||||||| | || ||| | |||| |:|||| |
Db   925 -QNLSRAEVEKRADELSFPLSVVEFLQGHIGIPHGGFPEPFRSKVLKSLPRIEGR-PGAS 982

Qy   950 VPEEEQAHLDA------DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALLDDREFFYG 1003
            |||:||| | |:|||:|||:||:||| |||||||:|:||||||||||:|:||
Db   983 LPPMDFEALESGLRAAHGDEITPEDVMSAAMYPKVFQEEFKEFTSTFGPVDCLNTRLFLDG 1042

Qy   1004 LVEGRETLIRLPDVRTPLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTAT 1063
            | |:| |||:| ||| |||:|||:|||| |:||| |:|| ||: |:||| ||| ||:|
Db   1043 PKIAEEFQVELERGKI-LHIKALALGDLNKSGQREVFFELNGQLRSVLVKDTAAMKEMHF 1101

Qy   1064 AEKADSSNKGHVRAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRV 1122
            || |:||| | | | | |:|||:|||:|||||||: | |:|:||:|:||::||:|||
Db   1102 HPKALKDVRGQVGAPMPGKVVEVKVKAGQKVERGQPLCVLSAMKMETVVNSPISGIISKV 1161

Qy   1123 VVPAATKVEGGDLIVVVS 1140
            :| :|| : || ||:| ::
Db   1162 HVNADSSLEGEDLILEIT 1179

```
RA   Kondo H., Kazuta Y., Saito A., Fuji K.-I.;
RT   "Cloning and nucleotide sequence of Bacillus stearothermophilus
RT   pyruvate carboxylase.";
RL   Gene 191:47-50(1997).
CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE CARBOXYL
CC       GROUP TO PYRUVATE IN THE SECOND. CATALYZES THE INITIAL REACTIONS
CC       OF GLUCOSE SYNTHESIS FROM PYRUVATE (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP + PYRUVATE + HCO(3)(-) = ADP +
CC       ORTHOPHOSPHATE + OXALOACETATE.
CC   -!- COFACTOR: BIOTIN (BY SIMILARITY).
CC   -!- PATHWAY: GLUCONEOGENESIS.
CC   -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, PARTICULARLY FROM
CC       EUKARYOTES, LIPOAMIDE TRANSFERASES AND CARBAMYL PHOSPHATE
CC       SYNTHETASES.
DR   EMBL; D83706; BAA12072.1; -.
DR   HSSP; P24182; 1DV1.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005481; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; FALSE_NEG.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Ligase; Multifunctional enzyme; Biotin; Gluconeogenesis; ATP-binding.
FT   DOMAIN       1     520       BIOTIN CARBOXYLASE (BY SIMILARITY).
FT   DOMAIN     521     968       CARBOXYLTRANSFERASE (BY SIMILARITY).
FT   DOMAIN    1064    1147       BIOTIN CARBOXYL CARRIER PROTEIN
FT                                (BY SIMILARITY).
FT   NP_BIND    167     172       ATP (BY SIMILARITY).
FT   ACT_SITE   296     296       BY SIMILARITY.
FT   BINDING   1112    1112       BIOTIN (BY SIMILARITY).
SQ   SEQUENCE   1147 AA;  128336 MW;  DFD2F003F838F591 CRC64;

Query Match             43.3%;  Score 2505;  DB 2;  Length 1147;
Best Local Similarity   46.7%;  Pred. No. 5e-126;
Matches  538;  Conservative 182;  Mismatches 397;  Indels 34;  Gaps 15;

Qy     13  KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL    72
           ::|||:||||||||:|: : :|::|||  ::|:|||:|||:||:|||||: :||:||||
Db      7  RKVLVANRGEIAIRVFRACTELGIRTVAIYSKEDVGSVHRYKADEAYLVGEGKKPIEAYL    66

Qy     73  DIDEIIGAKKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT   132
           ||: : |:| ||:|:||:|||||||:|| |::: ||||:|:||  ||:|  ||||:
Db     67  DIEGIIETAKAHDVDAIHPGYGFFSENIQFAKRCREEGIIFFGPNENHLDMFGDKVKARH   126

Qy    133  AAKKAGLEPVLAEST-PSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKLA  191
           ::|:: |:  |::   ::  |: |::|:| :|||:||||   ::::||||:|:||:|:::
Db    127  AAVNAGIPVIPGSDGPVDGLEDVVAFAEAHGYPIIIKALGGGGRGMRFVRSKSEVKEAF   186

Qy    192  TEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVEI   251
           |::|:|::|:|||:|||::: |:||||||||:||| ::||:|||||||:|::|:|||||:
Db    187  ERAKSEAKAAFGSDEVVVEKLIENPKHIEVQILGDYEGNIVHLYERDCSVQRRHQKVVEV   246

Qy    252  APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLV--DEKGNHVFIEMNPRIQVEHTV   309
           :|:: || ||::||| :: ::  ::|: ::|:|||  |||  :|||:||:|||||||:|
Db    247  APSVSLSDELRQRICEAAVQLMRSVGYVNAGTVEFLVSGDE---FYFIEVNPRIQVEHTI   303

Qy    310  TEEVTEVDLVKAQMRLAAGATL-KELGL-TQDKIKTHGAALQCRITTEDPNNGFRPDTG   366


Qy    355  EDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQLGGE-ITAHFDSMLVKMTCRGSDFETAV   413
           ||:::|||||| :||:: ::  |:|||||| :|| | ||:||||||:|:|::||:| ||
Db    387  EDPSKGFSPDTGKIEVYRSAGNGVRLDGGNGFAGAIITPHYDSMLVKVCTCRGSTYEIAR   446

Qy    414  ARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRI   473
           ||:|||| | ::||| |||| ||| |::|::||: :|||:||||||| |: || |:|||
Db    447  RKVVRALVEFRIRGVKTNIPFLTSLLSHPVFVDGTCWTTFIDDTPELFALVGSQNRAQKL   506

Qy    474  LDYLADVTVN-----KPHGVRPK---DVAAPI--DKLPNIKDLPLP--RGSRDRLKQLGPA  522
           | | |: :      ::: |  :   :|   |  |::  :|: : |  ||  :|||:
Db    507  LAYLGDVAVNGSSIKGQIGEPKLKGDIIKPVLHDAAGKPLDVSVPATKGWKQILDSEGPE   566

Qy    523  AFARDLREQDALAVTDFTERDAHQSLLATRVRSFALKPAAEFAVAKLTPELLSVEAWGGAT  582
           |||:: |:| :|| | : :|::||| ||| ::|: ||||: ||  :||: ||:|: ||:|
Db    567  AFARAVRANKGCLIMDTWRDAHQSLLATRVRTIDLLNIAHETSHALANAYSLECWGGAT   626

Qy    583  YDVAMRFLFEDPWDRLDELRERAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGV  642
           :|||||||||||||||||: :||:||||| |||||||| |  ::| ::: | :|||:|||
Db    627  FDVAMRFLYEDPWDRLRKLRKAVPNIPFQMLLRGANGVAYSSLPDNAIYHFCKQAKKCGV   686

Qy    643  DIFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYILKMAEEI   702
           |||| |||||| :: ::: :|| :|  ||| | ||::|:::||:|| | || :|:||||
Db    687  DIFRVFDALNDVDQLEVGIKAV-HAAEGVVEATICYSGDMLNPSKK-YNLPYLDLVDKV    744

Qy    703  VKSGAHILAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVHTHDTAGGQLATYFAAAQAG  761
           ||: | :||||||||  :  : ::|: ||: :  :||:|||||:|||:|:| |:   :|
Db    745  VQFKPHVLGIKDMAGVLKPQAARLLIGSIRERYPDLPIHVHTHDSAGTGVASMIACAQAG   804

Qy    762  ADAVDGASAPLSGTTSOPSLSAIVAAFAHTRDTGLSLEAVSDLEPYWEAVRGLYLPFES   821
           ||||: : : :|:    |:||| :|:   | : :|::  ::|:: | |: |:  :|: |
Db    805  ADAVDRATDSLSGMTSOPSIGAILASLEGTEHDPGLNSAQVRALDTYWAQLRLLYSPFEA   864

Qy    822  GTPGPTGRVYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVIPSSK   881
           ||:| ||| ::|:: |||:|| :   |: |  ::| ::::  :   :: :| ||||::|
Db    865  GLTGPDPEVYEHEIPGGQLTNLIFQASQLGLGQQWAETKKAYESANDLLGDVVKVIPTSK   924

Qy    882  VVGDLAIHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRS   941
           ||||| :|| | ||  :    || || :   :   |||::||  ::| |:|:  || |:
Db    925  VVGDLAQFMVSNKLTAEDVIARAGELDFPGSVLEFLEGLMGQPYGGFPEPLRSRALRDRR   984

Qy    942  E-GKAP---------LIEVPEEEQAHLDADDSKERRNSLNRLLFPKPTEEFLEHRRR     988
           : |:||         | : :|::  :  :   ::: :| ||::  ||::|:|::::
Db    985  KLDKRPGLYLEPLDLAKIKSQIRENYGAATEYDVA------SYAMYPKVFEDYKKFVAK  1037

Qy    989  FGNTSALDDREFYGLVEGRETLIRLPDVRIPLLVRLDAISE-PDDKGMRNVVANVNGQI  1047
           || |::|:: || | :  :|  :: :|::  |: :   :|: ||  ||:| |:  :|:
Db   1038  FGDLSVLPTRYFLAKPEIGEEFHVELEKGKV-LILKLLAIGPLSEQTGQREVFYEVNGEV 1096

Qy   1048  RPMRVRDRSVESVTATAEKADSSNKGHVAAPFAGVVT-VTVAEGDEVKAGDAVAIIEAMK 1106
           || :|:| |: |: |:| :   |:|||| | ||||  |||| | ||||| ||| :|||
Db   1097  RQVSVDDKKASVENTARPKAELGDSSQVGAPMSGVVVEIRVHDGLEVKKGDPIAVLSAMK 1156

Qy   1107  MEATITASVDGKIDRVVVPAATKVEGGDLI  1136
           || : |:  |  |  ||:|| :||| ||:
Db   1157  MEMVISAPHSGKVSSLLVKEGDSVDGQDLV  1186


RESULT 14
P94448
ID   P94448       PRELIMINARY;       PRT;   1147 AA.
AC   P94448;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyruvate carboxylase (EC 6.4.1.1) (Pyruvic carboxylase) (PCB).
OS   Bacillus stearothermophilus.
OC   Bacteria; Firmicutes; Bacillales; Geobacillus.
OX   NCBI_TaxID=1422;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K1041;
RX   MEDLINE=97354293; PubMed=9210587;
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=EGD-e / Serovar 1/2a;
RX   MEDLINE=21537279; PubMed=1679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
DR   EMBL; AL591977; CAC99150.1; -.
DR   Listlist; LM001072; -.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Complete proteome.
SQ   SEQUENCE 1146 AA;  128049 MW;  A062F88C9A092B6F CRC64;

Query Match        43.1%;
Best Local Similarity 45.8%;
Matches 526;   Conservative 185;   Mismatches 408;   Indels 30;   Gaps 11;

Qy    13  KKILVANRGEIAVRAFRAALETGAANTVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL  72
Db     5  KKVLVANRGEIAIRVMRACTELKIKTVAIYSQEDTGSFHRKSDEAYLVGAGKKPIDAYL  64

Qy    73  DIDEITGAAKKVKADIYPGVYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT  132
Db    65  DIENIIEIAKESGADAIHPGYGFLSENIEFARRCEQEGIIFVGPKSKHLDMFGDKIKAKE  124

Qy   133  AAKKAGLPVLAEST-PSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRFVASPDELRKLA  191
Db   125  QALLADIPVIPGSNGPVAGIKEVEEFGEKNGYPLMIKASLGGGGRGMRVVESKEHVKESF  184

Qy   192  TEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVEI  251
Db   185  ERASSEAKAAFGNDEVYVEKCVMNPKHIEVQILGDTHGNIVHLFERDCSIQRRHQKVEV  244

Qy   252  APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE  311
Db   245  APCNAITSELRNRICDAAVKLMKNVDYINAGTVEFLV-EGDDFYFIEVNPRVQVEHTITE  303

Qy   312  EVTEVDLVKAQMRLAAGATLKELGLIT--QDKIKTHGAALQCRITTEDPNNGFRPDTGTI  368
Db   304  MITGIDIVQSQLFIADGYALAHDQLVAIPKQEDIHIHGSAIQSRITTEDPLNNFMPDTGRV  363

Qy   369  TAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSG  427
Db   364  DTYRSTGGFGVRLDAGNGFQGTVVTPFYDSLLVKLCTWGMTFEQATRKMRRNLIEPRIRG  423

Qy   428  VATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVNKPHG  487
Db   424  VKTNIPFLLNVVRHPDFASGNYNTSFIDTTPELFKFPHIRDRGTKTLRYIGNVTVNGFPG  483
```

```
Db   304  TEMTGIDIVQSQILIADGCSLHSHEVGIPKQEDIRINGYAIQSRVTTEDPLNNFMPDTG  363
Qy   367  TITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTV  425

Db   364  KIMAYRSGGGFGVRLDAGNGFQGAVITPYDSLLVKLSTWALTFEQAARKMLRNLRTFRI  423

Qy   426  SGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVNKP  485

Qy   424  RVFKTNIPFLENVVQHPKFLSGEYDTSFIDTTPELFVFPRRKDRGTKMLTYIGITVTVNGF  483

Qy   486  HGVRPKDVAAPI--DKLPNIKDL-PLPRGSRDRIKQLGPAAFARDLREQDALAVTDTIF  541
Db   484  PGIGKK--KKPVFDKPRLPKLSEAEPIPAGTKQIIDKHGPEGLVRWIQEQPRVLLTDTIF  541

Qy   542  RDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDEL  601
Db   542  RDAHQSLLATRVRTVDLVRAAEPSARLLPNLFSLEMWGGATFDVAYRFLKEDPWDRLLKL  601

Qy   602  REAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAI  661
Db   602  RDAFPNVLFQMLLRSANAVGYKNYPDNVIREFVEKSAHAGIHVFRIFDSLNWVKGMTVAI  661

Qy   662  DAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYIKMAEEIVKSGAHILAIKDMAGLLRP  721
Db   662  DAVRQSG-KIAEAAICYTGDILDPSRSKYNLDYYKALAKELEQAGAHILAIKDMAGLLKP  720

Qy   722  AAVTKLVTALRREFDLPVHVTHHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSL  781
Db   721  QAAHVLISGLKETVDIPIYLHTHDTSGNGIYTYAKAIEAGVDIVDVAISSMAGLTSQPSA  780

Qy   782  SAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLS  841
Db   781  NTLIYALEGTERAPEVDIYGLEQLARYWEDVRKFYQEFESGMNAPHTEVYMHEMPGGQYS  840

Qy   842  NLRAQATAIGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFA  901
Db   841  NLQQQAKAVGLGDRWDEVKEMYRRVNDLFGDIVKVTPSSKVVGDMALYMVQNNLTEQDIF  900

Qy   902  ADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDAD  961
Db   901  ERGETLNFPDSVVELFEGYLGQPHGGFPEKLQRIILKGRE---PITVRPGELLEPVDFE  956

Qy   962  D-SKERRNSLNR-------LLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRET  1010

Db   957  QMKKELYDKLGREVTDFDAIAYAILPKVLEYAETVEKYGDVSVLDTPIFLYGMRLGEEI  1016

Qy   1011 LIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSS  1070

Db   1017 EVEIERGKT-LIVKLVSIGQPQADGTRVVYFELNGQPREVVIRDESIKTAVVEHIKADRT  1075

Qy   1071 NKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATK  1129

Db   1076 NPNHIAATMPGTVVKVLVEKGEKVDKGDHLMVTEAMKMETTVQAPFAGVVKDIYVKSGDA  1135

Qy   1130 VEGGDLIVVVS  1140
Db   1136 IQAGDLLMELS  1146
```

RESULT 15
Q8Y846
ID   Q8Y846;   PRELIMINARY;   PRT;   1146 AA.
AC   Q8Y846;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   PycA protein.
GN   PycA OR LMO1072.
OS   Listeria monocytogenes.
OC   Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX   NCBI_TaxID=1639;
RN   [1]

```
Qy   488  VRPKDVAAPIDKLPNIKDLP----LPRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRD  543
          ::|:  |: |:|:    :  ||:  :|  ||  :  :||||  |
Db   484  IKHRD--KPVYAEPRLPKIPYGSQISPGTKQILDAKGPEGVVDWVKKQEEVLLTDTTLRD  541

Qy   544  AHQSLLATRVRSFALKPAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELRE   603
          ||||||||||  ||:|:|  :|  | :| ||:||||  |:|  ||  ||: ||::
Db   542  AHQSLLATRVRSKDIFQVADAMAHLLPNMFSFEMWGGATFDVAYRFLNEDPWVRLETLRK   601

Qy   604  AMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDA   663
          :|||  ||||||  ||:|  ||||||||:   ||||||:|:|||||:||:|  : :|||
Db   602  QIPNVMFQMLLRGANAVGYKNYPDNVIREFVKQSAQSGVDVFRVFDSLNWIKGMEVSIDA   661

Qy   664  VLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAA   723
          |    |:: |:|:||::||  |:|||||:|  | |||:|||||||| | |||||||||:  |
Db   662  VREAG-KVVEAAICYTGDIDDDYRTKYTIDYYKDMAKELVAQGTHILGIKDMAGLLKPQA   720

Qy   724  VTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSA   783
          | :  || | |  | :| |  | ||:|| | ::| ||| |:|:: ||||||::
Db   721  AYRLIGELKDTVDPIHLHTHDTSGNGIYTYAAVSAGVDIVDVASSAMSGATSQPSMTG   780

Qy   784  IVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSNL   843
          |:| | |  : :| |: :| ||||  :::|   ||  | |||||:|  :  ||::||:
Db   781  LYYGLVNGNRQTNLDAQNSQIINHYWEDVRHYYKDFDNALNSPQTEVIIHEMPGGQYTNL   840

Qy   844  RAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVTPSSKVVGDLALHLVGAGVDPADFRAD   903
          : ||  |||:| :|:  :| :  :  :   |||||||||||||||||  :  :
Db   841  QQQAIAVGLGDRKWDEVKEMYTVVNQMFGDIVKVTPSSKVVTPSSKVVGDLALFMVQNELSEEDVYEK   900

Qy   904  PQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDADDS   963
          | |   ||:|  |:| |   |   :   : :  | :
Db   901  GDTIDFPDSVIEFFMGEIGQPYGGFPEKLQKLVLKGRT---PLTDRPGALMEPVNFVDV   956

Qy   964  K------ERRNSLNRLLFPKPTEEFLEHRRFGNTSALDDREFFYGLVEGRETLI   1012
          :       ::  ::: ||  ::|:|:: :  :  ::||: |:::|  |:|
Db   957  KAELKEKMGYEPTEKDVISYILYPKVFLDYQDMINKYGDVTVLDTPTFYKGMRLGETIEV   1016

Qy   1013 RLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNK   1072
          :|  :  : ||||:| | ||  || :|     | ::  :| : |:||:  :|
Db   1017 ELEKGKI-LLIKLNSIGEPIADGTRVIYFELNGQPREINIQDMNVQSTVIARRKIDTTNP   1075

Qy   1073 GHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVE   1131
          || :  : |  : |  | :  | |::| |  :| ::|  ::|  |: |: :|
Db   1076 EHVGATMTGSVIQVVKGDSVKKGDPLLITEAMKMETTIQAPFDGEVSSIYVSDGDTE   1135

Qy   1132 GGDLIVVS   1140
          || ||: |:
Db   1136 SGDLLIEVN   1144
```

Search completed: September 24, 2003, 15:59:16
Job time : 93.3413 secs

```
                    GenCore version 5.1.6
         Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:       September 24, 2003, 15:51:53 ; Search time 34.741 Seconds
                                  (without alignments)
                           3155.707 Million cell updates/sec

Title:          US-09-974-973A-19
Perfect score:  5788
Sequence:       1 MSTHTSSTLPAFKKILVANR........RVVVPAATKVEGGDLIVVVS 1140

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters:        283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      PIR_76:*
                1: pir1:*
                2: pir2:*
                3: pir3:*
                4: pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES

Result                %      Query
 No.     Score   Match Match Length DB  ID
-------  -----   ----- ----- ------ --  --
   1    3712.5   64.1  1127   2  D70671    pyruvate carboxyla
   2    2541.5   43.9  1178   2  A47255    pyruvate carboxyla
   3    2539.5   43.9  1178   2  JC4391    pyruvate carboxyla
   4    2525.5   43.6  1150   2  A83978    pyruvate carboxyla
   5    2524    43.6  1148   2  F69685    pyruvate carboxyla
   6    2517.5   43.5  1178   1  JC2460    pyruvate carboxyla
   7    2493    43.1  1146   2  AH1208    pyruvate carboxyla
   8    2491.5   43.0  1144   2  D97227    pyruvate carboxyla
   9    2490    43.0  1146   2  AC1565    pyruvate carboxyla
  10    2488    43.0  1150   2  G89881    pyruvate carboxyla
  11    2473.5   42.7  1174   2  AE2911    pyruvate carboxyla
  12    2473.5   42.7  1174   2  C97686    pyruvate carboxyla
  13    2464.5   42.6  1158   2  AE3285    pyruvate carboxyla
  14    2464    42.6  1175   2  T20346    pyruvate carboxyla
  15    2457    42.4  1178   1  QXBYP     pyruvate carboxyla
  16    2443    42.2  1185   2  T39734    pyruvate carboxyla
  17    2436    42.1  1180   2  S46094    pyruvate carboxyla
  18    2419    41.8  1195   2  T43735    pyruvate carboxyla
  19    2406    41.6  1137   2  E86708    pyruvate carboxyla
  20    2115.5   36.5   984   2  T44608    pyruvate carboxyla
  21    1079    18.6   501   2  D64453    biotin carboxylase
  22    1063.5   18.4   477   2  C70427    biotin carboxylase
  23    1036.5   17.9   447   2  A53311    biotin carboxylase
  24    1036.5   17.9   447   2  AH1923    biotin carboxylase
  25    1035    17.9   472   2  A70432    biotin carboxylase
  26    1019    17.6   506   2  D69277    biotin carboxylase
  27    997     17.2   491   2  A69123    biotin carboxylase
  28    981.5   17.0   448   2  S74380    biotin carboxylase
  29    976.5   16.9   471   2  G82966    probable biotin ca
```

```
  30    975.5   16.9  1095   2  B83471    probable pyruvate
  31    971.5   16.8  1078   2  D87647    hypothetical prote
  32    961.5   16.6   447   2  B97338    biotin carboxylase
  33    961     16.6   677   2  AC2997    hypothetical prote
  34    959     16.6   667   2  F98286    hypothetical prote
  35    957.5   16.5   444   2  C70444    biotin carboxylase
  36    957     16.5   449   2  AI0912    biotin carboxylase
  37    951.5   16.4   455   2  B86722    biotin carboxylase
  38    949     16.4   455   2  AD0445    biotin carboxylase
  39    948.5   16.4   539   2  T07093    acetyl-CoA carboxy
  40    946     16.3   444   2  T4481     biotin carboxylase
  41    946     16.3   448   1  F64105    biotin carboxylase
  42    942.5   16.3   455   2  C95049    acetyl-CoA carboxy
  43    942     16.3   450   2  A69581    acetyl-CoA carboxy
  44    941.5   16.3   455   2  A97920    biotin carboxylase
  45    939.5   16.2   457   2  H71553    probable biotin ca


                              ALIGNMENTS

RESULT 1
D70671
pyruvate carboxylase (EC 6.4.1.1) - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 11-Jan-2002
C;Accession: D70671; S73055
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
Rajandream, M.A.; Rogers, J.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
 Connor, R.; Davies, R.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.;
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genc
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70671
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1127 <COL>
A;Cross-references: GB:Z83018; GB:AL123456; NID:g3261671; PIDN:CAB05410.1; PID:g16948
A;Experimental source: strain H37Rv
R;Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, September 1994
A;Description: Mycobacterium tuberculosis cosmid tbc2.
A;Reference number: S73053
A;Accession: S73055
A;Molecule type: DNA
A;Residues: 1-353, 'TRAGSARCDPPAVPVSAWTAAPTWRRNQPVLRLHAGQADLS',396-1115,'EWRAETCWWW' <
A;Cross-references: EMBL:U00024; NID:g560506; PIDN:AAA50948.1; PID:g560527
C;Genetics:
A;Gene: pca; pyc
A;Start codon: GTG
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindi
C;Keywords: biotin binding; ligase; mitochondrion
F;4-457/Domain: biotin carboxylase homology <BCH>
F;1055-1127/Domain: lipoyl/biotin-binding homology <LPB>
F;1093/Binding site: biotin (Lys) (covalent) #status predicted

Query Match             64.1%;  Score 3712.5;  DB 2;  Length 1127;
Best Local Similarity   64.4%;  Pred. No. 3.4e-196;
Matches  730;  Conservative 153;  Mismatches 239;  Indels  11;  Gaps  6;

Qy     12  FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY  71
               |:|||||| |:||| ||:|| | ||||| |||||||||||||||||||||||||||||
Db      2  FSKVLVANRGEIAIRAFRAAAYELGVGTVAVYPYEDRNSQHRLKADESYQIGDIGHPVHAY  61

Qy     72  LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLLTGDKSRAV  131
           | |:||| |:||| ||||||||||||||||||||||||||||| |:|| ||||||||||
Db     62  LSVDEIVATARRAGADAIYPGYGFLSENPDLAAACAAAGISFVGPSAEVLELAGNKSRAI  121

Qy    132  TAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKLA  191
           |:||||||||||||:||| :|:|| :||:|||||||||||||||||||||||| |: :
Db    122  AAAREAGLPVLMSSAPSASVDELLSVAAGMFPLFVKAVAGGGGRGMRRVGDIAALPEAI  181
```

```
A:Title: Adipose pyruvate carboxylase: amino acid sequence and domain structure deduc
A:Reference number: A47255; MUID:93189578; PMID:8446588
A:Accession: A47255
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1178 <ZHA>
A:Cross-references: GB:L09192; NID:g293743; PIDN:AAA39737.1; PID:g293744
A:Experimental source: 3T3-L1 adipocytes
A:Note: sequence extracted from NCBI backbone (NCBIN:126874, NCBIP:126875)
C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindi
C:Keywords: biotin binding; ligase; mitochondrion
F:1-20/Domain: transit peptide (mitochondrion) #status predicted <TRP>
F:21-1178/Product: pyruvate carboxylase #status predicted <MAT>
F:39-494/Domain: biotin carboxylase homology <BCH>
F:1105-1178/Domain: lipoyl/biotin-binding homology <LPB>
F:1144/Binding site: biotin (Lys) (covalent) #status predicted

Query Match      43.9%;  Score 2541.5;  DB 1;  Length 1178;
Best Local Similarity 47.0%;  Pred. No. 9.4e-132;
Matches  537;  Conservative 187;  Mismatches 187;  Indels 401;  Gaps  11;
```

```
Qy    13 KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSPHRSFASEAVRIGTEGSPVKAYL  72
Db    38 KKYVMVANRGEIAIRVFRACTELGIRTVAVYSEQDTGQMHRQKADEAYLLGRGLAPVQAYL  97

Qy    73 DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT 132
Db    98 HIPDIIKVAKENGVDAVHPGYGFLSERADFAQACQDAGVRFIGPSPEVVRKMGDKVEARA 157

Qy   133 AAKKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKlA 191
Db   158 IAIAAGVPVVPGTDSPISSLHEAHEFSNTFGFPIIFKAAYGGGGRGMRVVHSYEELEENY 217

Qy   192 TEASREAEAAFPDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVEI 251
Db   218 TRAYSEALAAFGNGALFVEKFIEKPRHIEVQILGDQYGNILHLYERDCSIQRRHQKVVEI 277

Qy   252 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE 311
Db   278 APATHLDPQLRSRLTSDSVKLAKQVGYENAGTVEFLVDKHGKHYFIEVNSRLQVEHTVTE 337

Qy   312 EVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAY 371
Db   338 EITDVDLVHAQIHVSEGRSLPDLGLRQENIRINGCAIQCRVTTEDPARSFQPDTGRIEVF 397

Qy   372 RSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 430
Db   398 RSGEGMGIRLDNASAFQGAVISPHYDSLLVKVIAHGDHPTAATKMSRALAEFRVRGVKT 457

Qy   431 NIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVNKPHGVRP 490
Db   458 NIPFLQNVLNNQQFLAGTVDTQFIDENPELFQLRPAQNRAQKLLHYLGHVMVNGPTTPIP 517

Qy   491 KDVA-APID-KLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDAHQSL 548
Db   518 VNVSPSPVDPAVPVVPIGPPPAGFRDILLREGPEGFARAVRNHQGLLLMDTTFRDAHQSL 577

Qy   549 LATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNV 608
Db   578 LATRVRTHDLKKIAPYVAHNFNKLFSMENWGGATFDVAMRFLYECPWRRLQELRELIPNI 637

Qy   609 NIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETN 668
Db   638 PFQMLLRGANAVGYTNYPDNVVFKFCEVAKENGMDVFRVFDSLNYLPNMLLGMEAAGSAG 697

Qy   669 TAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLV 728
Db   698 -GVEAAISYTGDVADPSRTKYSLEYYMGLAEELVRAGTHILCIKDMAGLLKPAACMLV 756

Qy   729 TALRREF-DLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAA 787
Db   757 SSLRDRFPDLPLHTHDTSGAGVAAMLACAQAGADVVDVAVDSMSGMTSQPSMGALVAC 816
```

```
Qy   192 TEASREAEAFPDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVEI 251
Db   182 EAASREAESAFGDPTVYLEQAVINPRHIEVQILADNLGDVIHLYERDCSVQRRHQKVIEL 241

Qy   252 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE 311
Db   242 APAPHLDAELRYKMCVDAVAFARHIGYSCAGTVEFLLDERGEYVFIEMNPRVQVEHTVTE 301

Qy   312 EVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAY 371
Db   302 EITDVDLVASQLRIAAGETLEQLGLRQEDIAPHGAALQCRITTEDPANGFRPDTGRISAL 361

Qy   372 RSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATN 431
Db   362 RTAGGAGVRLDGSTNLGAEISPYFDSMLVKLTCRGRDLPTAVSRARRAIAEPRIRGVSTN 421

Qy   432 IGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVNKPHGVRPK 491
Db   422 IPFLQAVLDDPDFRAGRVTTSFIDERPQLITARASADRGTKILNFLADVTVNNPYGSRPS 481

Qy   492 DVAAPIDKLPNIKDLPL---PRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDAHQS 547
Db   482 TI-YPDDKLP--DLDLRAAPPAGSKQRLVKLGPEGFARWLRESAAVGVTDTTFRDAHQS 537

Qy   548 LLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPN 607
Db   538 LLATRVRTSGLSRVAPYLARTMPQLLSVECWGGATYDVALRFLKEDPWERLATLRAAMPN 597

Qy   608 VNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVLET 667
Db   598 ICLQMLLRGRNTVGYTPYPEIVTSAFVQEATATGIDIFRIFDALNNIESMRPAIDAVRET 657

Qy   668 NTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKL 727
Db   658 GSAIAEVAMCYTGDLIDPGEQLYTLDYYLKLAEQIYDAGAHVLAIKDMAGLLRPPAAQRL 717

Qy   728 VTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAA 787
Db   718 VSALRSRFDLPVHLHTHDTGGGQLASYVAAWHAGADAVDGAAAPLAGTTSQPALSSTVAA 777

Qy   788 FAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSNLRAQA 847
Db   778 AAHTEYDTGLSLSSAVCALEPYWEALRKVYAFFESGLPGPTGRVYHHEIPGGQLSNLRQQA 837

Qy   848 TALGIADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAADPQKY 907
Db   838 IALGLGDRFEEIEEAYAGADRVLGRLVKVTPTSKVVGDLALVGAGVSADEFASDPARF 897

Qy   908 DIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDADDSKERR 967
Db   898 GIPESVLGFLRGELGDPPGGWPEPLRTAALAGRGAAR-PTAQLAADDEIALSSVGAK-RQ 955

Qy   968 NSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDA 1027
Db   956 ATLNRLLFPSPTKEFNEHREAYGDTSQLSANQFFYGLRQGEHRVKL-ERGVELLIGLEA 1014

Qy  1028 ISEPDDKGMRRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAGVTVTV 1087
Db  1015 ISEPDERGMRTVMCILNGQLRPVLVRDRSIASAVPRAEKADRGNPGHIAAPFAGVVTVGV 1074

Qy  1088 AEGDEVKAGDAVAIIEAMKMEATIIPASVDGKIDRVVVPAATKVEGGDLIVVVS 1140
Db  1075 CVGERVGAGQTIATIEAMKMEAPIIPAVPAVGTVERVAVSDTAQVEGGDLLVVVS 1127
```

RESULT 2
A47255
Pyruvate carboxylase (EC 6.4.1.1) precursor [similarity] - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Feb-2002
C:Accession: A47255
R;Zhang, J.; Xia, W.L.; Brew, K.; Ahmad, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 1766-1770, 1993

```
F;1105-1178/Domain: lipoyl/biotin-binding homology <LPB>
F;1144/Binding site: biotin (Lys) (covalent) #status experimental

Query Match           43.9%;    Score 2539.5;  DB 2;   Length 1178;
Best Local Similarity 46.5%;    Pred. No. 1.2e-131;
Matches 542;  Conservative 185;  Mismatches 410;  Indels 29;  Gaps 12;

Qy    1 MSTHTSSLPA----------------FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRG 48
                            | |||    |||::|||||   |  |||||||| :  | :||
Db   14 LGVRRSSTAPVASPNVRRLEYKPIKKVMVMANRGEIAIRVFRACTELGIRTVAVYSEQDTG 73

Qy   49 SFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAE 108
            ::  : |  |  | || |  ||||  :  |||||| ||||:|||||||||  :||   :
Db   74 QMHRQKADEAYLIGRGLAPVQAYLHIPDIIKVKENGVDAVHPGYGFLSERADFAQACQD 133

Qy  109 NGITFIGPTPEVLLDTGDKSRAVTAAKKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIFV 167
            || ||||:|||: :| ||: ||  | ||| ||  | :||:|  :|  |||  |:|: |:
Db  134 AGVRFIGPSPEVVRKMGDKVEARAIAIAAGVPVVPGTNSPINSLHEAHEFSNTYGFPIIF 193

Qy  168 KAYAGGGGRGMRFVASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDH 227
            || ||||||||| | |::|::  ||||:|| ||||| ||:|| : | | |||||||||:
Db  194 KAAYGGGGRGMRVVHSYEELEENYTRAYSEALAAFGNGALFVEKFIEKPRHIEVQILGDQ 253

Qy  228 TGEVVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFL 287
            | : :||||||| :|||||||||||||: |||:||.|: :|:||  :| ||: |||||||
Db  254 YGNILHLYERDCSIQRRHQKVVEIAPATHLDPQLRSRLTSDSVKLAKQVGYENAGTVEFL 313

Qy  288 VDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAA 347
            || :|:| ||| | |:|||||||| :| ||||:||: : ||::| |||| |:  | || |
Db  314 VDKHGKHYFIEVNSRLQVEHTVTEITDVDLVHAQIHVSEGRSLPDLGLRQENIRINGCA 373

Qy  348 LQCRITTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAACL-GGEITAHFDSMLVKMTCRG 406
            :||:|||||| . | |||| | | ||. |:|:||| |:   |  |::|:||:|||: : |
Db  374 IQCRVTTEDPARSFQPDTGRIEVRSGEGMGIRLDNASAFQGAVISPHYDSLLVKVIAHG 433

Qy  407 SDFETAVARAQRALAEFTVSGVAINIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPA 466
              . |:| :.|:|||||.| ||:::| | :: || :: |:|||:||:|::| :|||  |
Db  434 KDHPTAATKMSRALAEFRVRGVKTNIPFLQNVLNNQQFLAGIVDTQFIDENPELFQLRPA 493

Qy  467 DDEQGRILDYLADVTVNKPHGVRPKDVA-APIDKL-PNIKDLPLPRGSRDRLKQLGPAAF 524
            :: : : | ||..|.|| | .:  ::|  |||: : ||:  |  ..| .:  | |||  |
Db  494 QNRAQKLLHYLGHVMVNGPTPIVFKVSPSPVDPIVPVVPIGPPPAGFRDILLREGPEGF 553

Qy  525 ARDLREQDALAVDTDTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYD 584
            |  :| :  |||||:|||||||||||||||: | |  |: || :|| ||||| |||||:|
Db  554 ARAVRNHQGLLLMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFNNLFSIENWGGATFD 613

Qy  585 VAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDI 644
            |||||| | ||.||:||||::||: :|||||| | |||| |||:| . | : |  :|:|:
Db  614 VAMRFLYECPWRRLQELRELIPNIPFQMLLRGANAVGYTNYPDNVVFKFCEVAKENGMDV 673

Qy  645 FRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLYKMAEEIVK 704
            ||||| || :  :  ::  : |:| ::  ||: |:||::||   .| |:|  :|||| |:
Db  674 FRIFDSLNYLPNMLLGMEEAGSAG-GVVEAAISYTGDVADPSRTKYSLEYYMGLAEELVR 732

Qy  705 SGAHILAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVHTHDTAGGQLATYFAAAQAGAD 763
             |:||| ||||||| |||.|.:|| |||:.| |||:|:|||||:| |:|  ||||||||:
Db  733 AGTHILCIKDMAGLLKPAACTMLVSSLRDRFPDLPLHIHTHDTSGSGVAAMLACAQAGAD 792

Qy  764 AVDGASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGT 823
            .|| :: | ||||||||:|::||: :||:::|.|||||::|:||||| |||||: || .|
Db  793 VVDVAVDSMSGMTSQPSMGALVACTKGTPLDTEVPLERVFDYSEYWEGARGLYAAFDCTA 852

Qy  824 PGPTGR--VYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSK 881
                .|  ||::||||||:::|:.||: :||| :|:|::::|| ||::|| .:||||||
Db  853 TMKSGNSDVYENEIPGGQYTNLHFQAHSMGLGSKFKEVKKAYVEANQMLGDLIKVTPSSK 912

Qy  882 VVGDLALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRAL----- 937
            :||||| :   .| .   | | |:| | | ||: ||:| :|:| || |||:|::|
Db  913 IVGDLAQFMVQNGLSRAEAEAQAELSFPRSVVEFLQGYIGIPHGGFPEPFRSKVLKDLP 972

Qy  938 -EGRSEGKAPLTEVPEEEQAHLDA-DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSA 994
```

```
Qy  788 FAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGR--VYRHEIPGGQLSNLRA 845
            | : | |.|||||.:| |||||| |||||: || .|     |  ||::||||||::|:
Db  817 TKGTPLDTEVPLERVFDYSEYWEGARGLYAAFDCTATMKSGNSDVYENEIPGGQYTNLHF 876

Qy  846 QATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAADPQ 905
            ||: :||| :|:|:::.|| ||::|| .:||||||::|||| :   .|  .    | | |
Db  877 QAHSMGLGSKFKEVKKAYVEANQMLGDLIKVTPSSKIVGDLIKVTPSSK 936

Qy  906 KYDIPDSVIAFLRGELGNPPGGWPEPLRTRAL-----EGRSEGKAPLTEVPEEEQAHLD 959
            : | | ||: ||:| :|.| || |||:|::|       |||:||| ||||||| |.| |
Db  937 ELSFPRSVVEFLQGYIGIPHGGFPEPFRSKVLKDLPRIEGRPGASLPPLNKELEKDLID 996

Qy  960 A-DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVR 1018
              .: .| |:: | | .     | |  |  || |:||::| | | ||||||||||||||:
Db  997 RHGEEVTPEDVLSAAMYPDVFAQFKDFTATFGPLDSLNTRLFLQGPKIAEEFEVELERGK 1056

Qy 1019 TPLLVRLDLAISEPDDKGSMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAP 1078
             | ::: | ||: | || .  :.  ::|| ||  .  :     :.::| : :..   | .|
Db 1057 T-LHIKALAVSDLNRAGQRQVFFELNGQLRSIIVKDTQAMKEMHFHPKALKDVKGQIGAP 1115

Qy 1079 FAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIV 1137
            | . ::.: |:.||.|.||  : :::|||||:::|:| ::||: :.:  : |:||.|||:
Db 1116 MPGKVIDIKVAAGDKVAAGQPLCVLSAMKMETVVTSPMEGTIRKVHVTKDMTLEGDDLIL 1175

Qy 1138 VV 1139
            :|
Db 1176 EI 1177
```

RESULT 3
JC4391
Pyruvate carboxylase (EC 6.4.1.1) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Jan-1996 #sequence_revision 25-Apr-1997 #text_change 11-Jan-2002
C;Accession: S68252; S72393; JC4391; S06440
R;Jitrapakdee, S.; Booker, G.W.; Cassady, A.I.; Wallace, J.C.
Biochem. J. 316, 631-637, 1996
A;Title: Cloning, sequencing and expression of rat liver pyruvate carboxylase.
A;Reference number: S68252; MUID:96257760; PMID:8687410
A;Accession: S68252
A;Molecule type: mRNA
A;Residues: 1-178 <JIT1>
A;Cross-references: EMBL:U36585; NID:g1040973; PIDN:AAC52668.1; PID:g1040974
A;Accession: S72393
A;Molecule type: protein
A;Residues: 489-505 <JIT2>
A;Experimental source: liver
R;Lehn, D.A.; Moran, S.M.; MacDonald, M.J.
Gene 165, 331-332, 1995
A;Title: The sequence of the rat pyruvate carboxylase-encoding cDNA.
A;Reference number: JC4391; MUID:96096548; PMID:8522203
A;Accession: JC4391
A;Molecule type: mRNA
A;Residues: 1-221,'P',223-865,'D',867-976,'G',978-1178 <LEH>
A;Cross-references: GB:U32314; NID:g929987; PIDN:AAA96256.1; PID:g929988
A;Experimental source: liver
R;Thampy, K.G.; Huang, W.Y.; Wakil, S.J.
Arch. Biochem. Biophys. 266, 270-276, 1988
A;Title: A rapid purification method for rat liver pyruvate carboxylase and amino acid s
A;Reference number: S06440; MUID:89024676; PMID:3178228
A;Accession: S06440
A;Molecule type: protein
A;Residues: 'SG',23-25,'PL',28-29,'LL',32-34,'P',1134,'A',1136-1137,1139-1152,'T',1154-1
C;Genetics:
A;Gene: pc
A;Genome: nuclear
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding
C;Keywords: biotin binding; gluconeogenesis; homotetramer; ligase; mitochondrion
C;Comment: This enzyme is located in the mitochondrial matrix and catalyzes the conversi
intermediates that exit the mitochondrion for consumption in various pathways.
F;1-20/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F;21-178/Product: Pyruvate carboxylase #status predicted <MAT>
F;39-494/Domain: biotin carboxylase homology <BCH>

```
Db   484  GYPGLE-KTKKPVFDKPPVPKLKLSEPIPDGTKQILDQHGPEGLAKWVKEQKHVLLTDTT  542
                ||| :||  ||| : |:| ||  | ::|||
Qy   541  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  600
          ||||||||| :| |  || : ::| | :|: |:  :||||||
Db   543  FRDAHQSLLATRVRTHDLKQIAEPTARLLPNLFSMEMWGGATFDVAMRFLHEDPWERLLI  602

Qy   601  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  660
          |::| || |:| |  :| :  | | | |::|||: :||||| : |
Db   603  LRKKAPNVLFQMLLRASNAVGYKNYPDNLIREFVDKSANAGIDVFRIFDSLNWVEGMKLA  662

Qy   661  IDAVLETNTAVAEVAMAYSGDLSDPNEKILYTLDYLKMAEEIVKSGAHILAIKDMAGLLR  720
          : |::| ::| |  | :|| :::: :: || :|| |:|||| : || |:| ||||||:
Db   663  IEAVGEAN-KIAEATICYTGDILDSSRPKYDLAYYKKLAKELEAAGAHILGIKDMAGLLK  721

Qy   721  PAAVTKLVTALRLRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  780
          | |  |:| || |:|||| :|::|:| |||:| :||| :||  |||:|| : : : ::
Db   722  PEAAYQLVAELKDTVTIPVHLHTHDTSGNGIFTYARALEAGVDIVDVAVSSMAGLTSQPS  781

Qy   781  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGVTGRVYRHEIPGGQL  840
          |:| | | :|| | |: :   ::    |  :|| :| :||:||   | | : :|  |
Db   782  ANSLYYALADSERQPNVNTALEQLAEFWEETRKFYAGFESGMNAPHTEVYEHEMPGGQY  841

Qy   841  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVIPSSKVVGDLALHLVGAGVDPADF  900
          |||:||| :|| |:|::   :|  ||:|  || ||:|||||||||:||:||  | |:
Db   842  SNLQQQAKAVGLGHRWNEVKKMRTVNDMFGDVVKVTPSSKVVGDMALYMVQNDLTEEV  901

Qy   901  AADPQKYDIPDSVIAFLRGELGNPPGGWEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA  960
          : |::|  : : | :|||::||  ||: ||| :   ||      :  : :|
Db   902  YENGHKLLDFPDSVVEFFGQLGQPYQGFPKKLQEIILKGRK---PITNRPGENMEPIQF  957

Qy   961  DDSKER-RNSLNR--------LLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRE  1009
          : :|:     :|:         ||   :|||  :| ||||  | |::||||  || |
Db   958  EAIKEELYNKLDRQVTSHDILSYALYPKVFMEFERRFRQTFGDVSVLDTPTFFYGLRPGEE  1017

Qy  1010  TLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVANVNGQIRPMRVRDRSVESVTATAEKADS  1069
          | :::| |||||| |||||||:|::||:|||::||  ||:||:| ||||||:|||||| ::
Db  1018  IEVEIEQGKT-LIVKFISLSKPQDDGNRIVYFFELNGQPREVLIKDQSVKTSIISRPKADX  1076

Qy  1070  SNKGHVAAPFAG--VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAAT  1128
          |||  ||||| |  |:|| |:| :||||||| ||||:|||||||:| |:|||||| ||:
Db  1077  SNPNHIGASMPGTVVKALVEKGDKVKQGDHLMITEAMKMETTVQAPFDGEVVALHVKVGD  1136

Qy  1129  KVEGGDLIVVV  1139
          |:|| ||:||:
Db  1137  AIQTGDLLIEV  1147
```

```
RESULT 5
F69685
pyruvate carboxylase (EC 6.4.1.1) pycA [similarity] - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 18-Aug-2000
C;Accession: F69685
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: F69685
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1148 <KUN>
```

```
Db   973  RIEGRPGASLPLNLKELEKDLIDRHGEEVIPEDVLSAAMYPDVFAQFKDFTATFGPLDS  1032
                :||| |:|| ::| |:::| | : || |: :|||: ||||| : :| || |:
Qy   995  LDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRD  1054
          :   || | |  ||  : |:| ::| ||: :|| :  :|  | : || |  | |: |
Db  1033  LNTRLFLQGPKIAEEFEVELERGKT-LHIKALAVSDLNRAGQRQVFFELNGQLRSILVKD  1091

Qy  1055  RSVESVTATAEKADSSNKGHVAAPFAG--VVTVTVAEGDEVKAGDAVAIIEAMKMEATITA  1113
          |:  ::|| :: ||  ::|| ||| |   |:|| |:|  |||||:| |:|||||| :| :
Db  1092  TQAMKEMHFHPKALKDVKGQIGAPMPGKVIDVKVAAGAKVVKGQPLCVLSAMKMETVVTS  1151

Qy  1114  SVDGKIDRVVPAATKVEGGDLIVVV  1139
          | : |:|:|||: | ||| || :|:
Db  1152  PMEGTIRKVHVTKDMTLEGDDLILEI  1177
```

```
RESULT 4
A83978
pyruvate carboxylase pycA [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: A83978
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: A83978
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1150 <STO>
A;Cross-references: GB:AP001516; GB:AP001516; NID:gi10175192; PIDN:BAB06344.1; GSPDB:GN0(
C;Genetics:
A;Experimental source: strain C-125
C;Gene: PycA
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

Query Match     43.6%;  Score 2525.5;  DB 2;  Length 1150;
Best Local Similarity 46.8%;  Pred. No. 6.9e-131;
Matches 539;  Conservative 185;  Mismatches 400;  Indels 27;  Gaps 13;
```

```
Qy     9  LPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPV  68
          | | ||||||||||||| ||| | |  | ||:|| :  :|: |  | : || | | |||
Db     4  LKNIKKVLVANRGEIAIRIFRACTELHIRTVAIYSKEDTGAYHRYKADEAYLVGEGKKPI  63

Qy    69  KAYLDIDEIIGAAKKVKADAIYPGVGYFLSENAQLARECAENGITFIGPTPEVLDLTGDKS  128
          :|||: ||:| ||: | |:| ||:|||||||: | ||  |||||||||| |||  |||:
Db    64  EAYLDIEGIIEIAKRHGVDAIHPGYGFLSENIEFAKRCHEEGIIFIGPELEHLVMFGDKV  123

Qy   129  RAVTAAKKAGLPVLAEST-PSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDEL  187
          :|::|    |||:: :    |  | :|:|  | | |||||:|| ||||||||:|:  |:|
Db   124  QAREQAIKANLPVIPGSDGPVSSLEDVKAFADKHGYPFIIKAALGGGGRGMRIVRSENDV  183

Qy   188  RKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQK  247
          | |::||::|:||| || || ::  |:|:| |||||||:|||||||||||||  ||||||
Db   184  QESYERAKSEAKAAFGNDEVYVEKFIENPKHIEVQILADKHGNTLHLYERDCSVQRRHQK  243

Qy   248  VVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEH  307
          ||| | ||| | ::|:||| |:| | ||||  || ||||||:  |:||:||||||||||
Db   244  VVEVAPSVSLSEDVRERICQAAVQLAENVNYVNAGTVEFLVDREGNFYFIEVNPRIQVEH  303

Qy   308  TVTEEVTEVDLVKAQMRLAAGATL--KELGL-TQDKIKTHGAALQCRITTEDPNNGFPRPD  364
          ||||| || ||||:   |:| |:     | |:  |:|||  || | || ||||||| | |
Db   304  TITEMVTGIDIVQSQLFIADGEHLHGDRLGIPKQEEIVCHGYAIQSRVTTEDPSNGFLPD  363

Qy   365  TGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEF  423
          ||:|:|||| | || ||   |  |:| ::|: || ::  |  | | | |   | :|  |
Db   364  TGRINAYRSGGGFGVRLDAGNGFQGAVITPYYDSLLVKVSTWALITPEGAAAKMLRNLREF  423

Qy   424  TVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVN  483
          | |: :|| |||  ||:  : : | :   | :||||| ||| :|:|  || | :  ||:
Db   424  RIRGIKTNIAFLENVVQHRQFLSGEVNTSFIDQTPELFVPFKRKDRGTKMLSFIGETIVN  483

Qy   484  KPHGVRPKDVAAPIDK--LPNIK-DLPLRGSRDRLKQLGPAAFARDLREQDALAVTDTT  540
          || |  ||: | |:|    | ||  ||  |||| |  ||| : | : || |   |
```

A;Cross-references: GB:Z99111; GB:AL009126; NID:g2633699; PIDN:CAB13359.1; PID:g2633857
A;Experimental source: strain 168
C;Genetics:
A;Gene: pycA
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding
C;Keywords: ligase
F;8-465/Domain: biotin carboxylase homology <BCH>
F;1073-146/Domain: lipoyl/biotin-binding homology <LPB>
F;1112/Binding site: biotin (Lys) (covalent) #status predicted

Query Match   43.6%;   Score 2524; DB 2;   Length 1148;
Best Local Similarity   46.0%;   Pred. No. 8.3e-131;
Matches   533;   Conservative 186;   Mismatches 392;   Indels   48;   Gaps   13;

Qy   11 AFKKIVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAAVRIGTEGSPVKA   70
          : :||:||||||| || :|| |||| | || ||:|: | ||:|::| |::| |||:|
Db   5 SIQKVLVANRGEIAIRIFRACTELNIRTVAYSKEDSGSYHRYKADEAYLVGEGKKPIDA   64

Qy   71 YLDIDELIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRA   130
          |||: ||| ||:| || |||||||| | || :|||| || | ||| | | : :
Db   65 YLDIEGIIDIAKRNKVDAIHPGYGFLSENIHFARRCEEEGIVFIGPKSEHLDMFGDKVKA   124

Qy   131 VTAAKKAGLPVLAEST-PSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRK   189
          || ||| |||:| : |:|| | ||:| || |||||||||| |||||||:| |: | |
Db   125 REQAEKAGIPVIPGSDGPAETLEAVEQFGQANGYPIIIKASLGGGGRGMRIVRSESEVKE   184

Qy   190 LATEASREAEAAFGDGAVYVERAVINPQHIEVQIILGDHTGEVVHLYERDCSLQRRHQKVV   249
          |: : ||| ||||||||||||:|||| ||||| ||||:| || ||| |||::||||:|
Db   185 AYERAKSEAKAAFGNDEVYVEKLIENPKHIEVQVIGDKQGNVVHLFERDCSVQRRHQKVI   244

Qy   250 EIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTV   309
          | |||||| |||| :| | ||| |: || ||||:|||||| |||||||||||||:|
Db   245 EVAPSVSLSPLRDQICEAAVALAKNVNYINAGTVEFLV-ANNEFYFIEVNPRVQVEHTI   303

Qy   310 TEEVTEVDLVKAQMRLAAGATL-KELGLLTQDK-IKTHGAALQCRITTEDPNNGFRPDTG   366
          :/: / | |: ::|:: | | | :|: | | : ||||: | :|||||| |||||
Db   304 TEMIGVDIVQTQILVAQGHSLHSKKVNIPEQKDIFTIGVAIQSRVTTEDPQNDFMPDTG   363

Qy   367 TITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTV   425
          ||||:|| | |||| | || |||||| | ::|||| | || :| :|:| :|||:|
Db   364 KIMAYRSGGGFGVRLDTGNSFQGAVITPYYDSLLVKLSTWALITFEQAAAKMVRNLQEFRI   423

Qy   426 SGVATNIGFLRALLREEDTSKRIATGTFIADHPHLLQAPPADDEQGRILDYLADVTVN--   483
          || ||:|| | : :: | || : | | | ||: |::| :|||:| ::||
Db   424 RGIKTNIPFLENVAKHEKFLTGQYPTSFIDTTPELFNFPKQKDRGTKMLTYIGNVTVNGF   483

Qy   484 ------KPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAV   536
          |||| |||||: | ||:|| :|::|::|:| ||::||| ||:| :|:|:|
Db   484 PGIGKKEKPAFDKPLGVKVDVDQQP-------ARGTKQILDEKGAEGLANWVKEQKSVLL   536

Qy   537 TDTTPRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLEDPWD   596
          ||||||||||||||||||| | |:|:|| |:||||| || ||||||||:||||:| |
Db   537 TDTTPRDAHQSLLATRIRSHDLKKIANPTAALWPELFSMEMWGGATFDVAYRFLKEDPWK   596

Qy   597 RLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQ   656
          |||::|: |||| |||:|||||||||| || :: |:||:| |||||||||:| :: :|
Db   597 RLEDLRKEVPNTLFQMLLRSSNAVGYTNYPDNVIKEFVKQSAQSGIDVFRIFDSLNWVKG   656

Qy   657 MRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYIKMAEEIVKSGAHILAIKDMA   716
          :||| ||::|| ||:|||: |:| | | ||||| ||:|| |:|||||: ||||:|
Db   657 MTLAIDAVRDTG-KVAEAAICYTGDILDKNRTKYDLAYYTSMAKELEAAGAHILGIKDMA   715

Qy   717 GLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTT   776
          ||| | |:| | |:|| ::| |::|||:|||| || :|: |||: | ||:||
Db   716 GLLKPQAAYELVSALKETIDIPVHLRTHDTSGNGIYMYAKAVEAGVDIIDVAVSSMAGLT   775

Qy   777 SQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGGTGRVYRHEIP   836
          |||: :|:| ||: |:| :| | :|: |||::|||: | ||||: | ||:: | |
Db   776 SQPSASGFYHAMEGNDRRPEMNVQGVELLSQYWSVRKYYSEFEGMKSPHTEIYEHEMP   835

Qy   837 GGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVD   896
          ||| | ||| ||:|::|| |:|:||| |||:| ||||||||||:|||||: || |
Db   836 GGQYSNLQQAKGVGLGDRWNEVKEMYRVNDMFGDIVKVTPSSKVVGDMALYMVQNNLT   895

Qy   897 PADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEE--- 953
          ||:|| ||:|||:|||||||||||||||||||||||| ||||| |:|||||:|||||
Db   896 EKDVYEKGESLDPDSVVELFKGNIGQPHGGFPEPLKIQKLILKGQE----PITVRPGELLE   951

Qy   954 ------------EQAHLDADDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFF 1001
                       :| |:|:|:|: | ||: | | |: |:| |::||:|||:|
Db   952 PVSFEAIKQEFKEQHNLEISD----QDAVAYALYPKVFTDYVKTTESYGDISVLDTPTFF 1007

Qy 1002 YGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVT 1061
          |||: |:| ::| | |::|| ||:|: : | |:: : ||:|::|| |:|::|:|
Db 1008 YGMTLGEIEVFIERGKT-LIVKLISIGEPQPDATRVVYFELNGQPREVIKDESIKSSV 1066

Qy 1062 ATAEKADSSNKGHVAAPFAGVVTVTVAE-GDEVKAGDAVAIIEAMKMEATITASVDGKID 1120
          :| | |||:| | :| :| |||||: ||:| ||| | | |||||:| | :: ||
Db 1067 QERLKADRTNPSHIAASMPGTVKVLAEAGTKVNKGDHLMINEAMKMETTVQAPFSGTIK 1126

Qy 1121 RVVVPAATKVEGGDLIVVV 1139
          ||| | ||||| ||::|:
Db 1127 QVHVKNGEPIQTGDLLLEI 1145

RESULT 6
JC2460
pyruvate carboxylase (EC 6.4.1.1) precursor - human
N;Alternate names: pyruvate:carbon dioxide ligase (ADP-forming)
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 18-Aug-2000 #text_change 01-Feb-2002
R;Accession: G01933; JC2460; B27883; S01469
R;Walker, M.E.; Jitrapakdee, S.; Val, D.L.; Wallace, J.C.
submitted to the EMBL Data Library, July 1995
A;Reference number: H00708
A;Accession: G01933
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1178 <WAL>
A;Cross-references: EMBL:U30891; NID:g1101028; PIDN:AAA82937.1; PID:g1101029
R;MacKay, N.; Rigat, B.; Douglas, C.; Chen, H.S.; Robinson, B.H.
Biochem. Biophys. Res. Commun. 202, 1009-1014, 1994
A;Title: cDNA cloning of human kidney pyruvate carboxylase.
A;Reference number: JC2460; MUID:94324922; PMID:8048912
A;Accession: JC2460
A;Molecule type: mRNA
A;Residues: 1-224, 'WP', 227-351, 'A', 353-384, 'PT', 387-485, 'DV', 488-637, 'R', 639-728, 'A',
A;Cross-references: GB:S72370; NID:g632807; PIDN:AAB31500.1; PID:g632808
R;Lamhonwah, A.M.; Quan, F.; Gravel, R.A.
Arch. Biochem. Biophys. 254, 631-636, 1987
A;Title: Sequence around the biotin-binding site of human propionyl-CoA carb
A;Reference number: A27883; MUID:87212051; PMID:3355348
A;Accession: B27883
A;Molecule type: mRNA
A;Residues: 1083-1178 <LAM>
A;Cross-references: GB:M26122; NID:g189657; PIDN:AAA36423.1; PID:g387003
R;Freytag, S.O.; Collier, K.J.
J. Biol. Chem. 259, 12831-12837, 1984
A;Title: Molecular cloning of a cDNA for human pyruvate carboxylase. Structural relat
A;Reference number: S01469; MUID:85030380; PMID:6548474
A;Accession: S01469
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1135-1178 <FRE>
A;Cross-references: EMBL:K02282; NID:g189655; PIDN:AAA60033.1; PID:g189656
C;Genetics:
A;Gene: GDB:PC
A;Cross-references: GDB:119472; OMIM:266150
A;Map position: 11q11-11q13.1
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindi
C;Keywords: biotin binding; gluconeogenesis; ligase; mitochondrion
F;1-20/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F;2-1178/Product: pyruvate carboxylase #status predicted <MAT>
F;39-494/Domain: biotin carboxylase homology <BCH>
F;1105-1178/Domain: lipoyl/biotin-binding homology <LPB>
F;1144/Binding site: biotin (Lys) (covalent) #status predicted

```
Query Match          43.5%;     Score 2517.5;  DB 1;   Length 1178;
Best Local Similarity 46.2%;    Pred. No. 2e-130;
Matches  536;  Conservative 184;  Mismatches 412;  Indels 29;  Gaps 12;

Qy    6  SSTLPA---------FKKILVANRGEIAVRAFRAALETGAANTVAIYPREDRGSFHRS   53
            :|| ||        :|||||||| ||  :| |:||  |||  |
Db   19  TSTAPAASPNVRRLEYKPIKKVMVANRGEIAIRVFRACTELGIRTVAIYSEQDTGQMHRQ   78

Qy   54  FASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITF  113
         :| ||  ||: |||:||| ||::|:|  |:|||||:|||||||| :|::: | | :|: |
Db   79  KADEAYLIGRGLAPVQAYLHIPDIIKVAKENNDAVHPGYGFLSERADFAQACQDAGVRF   138

Qy  114  IGPTPEVLDLTGDKSRAVTAAKKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIFVKAVAG  172
         |||:|||  : ||| :||  | ||:||:| :: :| :: ::::: ::  : |:|:|||| |
Db  139  IGPSPEVVRKMGDKVEARAIAIAAGVPVVPGTDAPITSLHEAHEFSNTYGFPIIFKAAYG  198

Qy  173  GGGRGMRFVASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVV  232
         ||||||| |    |  | : ||  :|| ||| ||: |:: :::| :||||||||: :|:
Db  199  GGGRGMRVVHSYEELEENYTRAYSEALAAFGNGALFVEKFIEKPRHIEVQILGDQYGNIL  258

Qy  233  HLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKG  292
         |||||||::|||||||||||||:|:|| ||:|: :|:||: : ||:::|||||||||:.|
Db  259  HLYERDCSIQRRHQKVVEIAPAAHIDPQLRTRLTSDSVKLAKQVGYENAGTVEFLVDRHG  318

Qy  293  NHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRI  352
           :|||:| ||:|||||||| | :|||:||: : ||:| |:|||:|::|:  | | |||:
Db  319  KHYFIEVNSRLQVEHTVTEEITDVDLVHAQIHVSEGRSLPDLGLRQENIRINGCAIQCRV  378

Qy  353  TTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQL-GGEITAHFDSMLVKMTCRGSDFET  411
         ||||| . | ||||:| .:||  |  :|||:.|:  ||::: :.||:|:|: : | |: :
Db  379  TTEDPARSFQPDTGRIEVFRSGEGMGIRLDNASAFQGAVISPHYDSLLYKVIAHGKDHPT  438

Qy  412  AVARAQARALAEFTVSGVATNIGFLRALIREEDTSKRIATGFIADHPHLLQAPPADDEQG  471
         | : :  ||||| .| ||:||| || :::  ::    || | | :  : |:  | ::
Db  439  AATKMSRALAEFRVRGVKTNIAFLQNVLNNQQFLAGTVDTQFIDENPELFQLRPAQNRAQ  498

Qy  472  RILDYLADVTVNKPHGVRP-KDVAAPIDK-LPNIKDLPLPRGSRDRLKQLGPAAFARDLR  529
          :|:||  | ||    ::  | : : : | :| : .    | |  |::  |  |||| :
Db  499  KLLHYLGHVMVNGPTTPIPVKASPSPTDFVVPAVPIGPPAGFRDIL.REGPEGFARAVR   558

Qy  530  EQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRF  589
          : | :  ||||||||||||||||:  || .. :|   .|||||||.||||||:|||||
Db  559  NHPGLLLMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFSKLFSMENWGGATFDVAMRF  618

Qy  590  LFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFD  649
         |:: || ||.||||::|:| :|||||| |.|||: ||::|   |.| | ..|:|:||:||
Db  619  LYECPWRRLQELRELIPNIPFQMLLRGANAVGYTNYPDNVVFKFCEVAKENGMDVFRVFD  678

Qy  650  ALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHI  709
         :||:: .   |: | |.|.. |.|.|:||||::.||.. ||:| ||:.:|||:|::|:||
Db  679  SLNYLPNMLLGMEAAGSAG-GVVEAAISYTGDVADPSRTKYSLQYYMGLAEELVRAGTHI  737

Qy  710  LAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVHTHDTAGGQLATYFAAAQAGADAVDGA  768
         | |||||||:| :. | |||:||::| ||| |:|||||:.||:|:: :|||||||.||.|
Db  738  LCIKDMAGLLKPTACTMLVSSLRDRFPDLPLHIHTHDTSGAGVAAMLACAQAGADVVDVA  797

Qy  769  SAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLLPFESGTPGPTG  828
         : : ||:||||::.|:||  ||||.||  :|| | |:|||:|.|||||   |: | | :.|
Db  798  ADSMSGMTSQPSMGALVACTRGTPLDTEVPMERVFDYSEYWEGARGLYAAFDCTATMKSG   857

Qy  829  R--VYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDL  886
             :|::|||||| :||: .|  :||.:|| ::::||||:|:|||  .|:||||||:|||
Db  858  NSDVYENEIPGGQYTNLHFQAHSMGLGSKFKEVKAAYVEANQMLGDLIKVTPSSKIVGDL  917

Qy  887  ALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRAL------EGR  940
         | |:| .|::. .: | .|:. :| || .|| || ||| |||:|| | ::      ||
Db  918  AQFMVQNGLSRAEAEAQAEELSFPRSVVEFLQGYIGVPHGGFPEPFRSKVLKDLPRVEGR  977

Qy  941  SEGKAPLTEVPEEEQAHLDA-DDSKERRNSLNRLLFPKTEEFLEHRRRFGNTSALDDRE  999
         ..  :|| ::  ::  ::.  ||.||||||:||||| |||||||||||||||::|| | |
Db  978  PGASLPPLDLQALEKELVDRHGEVTPEDVLSAAMYPDVFAHFKDFTATFGPLDSLNTRL 1037

Qy 1000  FFYGLVEGRETLIRLPDVRTLPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVES 1059
         | .|.: :. :|: : :| |:  ::|  |:  .|::.: :|  . ||:::| |:||:
Db 1038  FLQGPKIAEEFEVELERGKT-LHIKALAVSDLNRAGQRQVFFELNGQLRSILVKDTQAMK 1096

Qy 1060  VTATAEKADSSNKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGK 1118
           | :::| . :.:  ..|: | .::::.|.|.:|:|||: | : |||||:.|.|:: :|
Db 1097  EMHFHPKALKDVKGQIGAPMPGKVIDIKVVAGAKVAKGQPLCVLSAMKMETVTSPMEGT  1156

Qy 1119  IDRVVVPAATKVEGGDLIVVV 1139
         : :| | |: :||||||::::
Db 1157  VRKVHVTKDMTLEGDDLLLEI 1177
```

RESULT 7

AH1208

pyruvate carboxylase homolog pycA [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.
science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1208
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1146 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC99150.1; PID:g16410474; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: pycA
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindi

```
Query Match          43.1%;     Score 2493;   DB 2;   Length 1146;
Best Local Similarity 45.8%;    Pred. No. 4.2e-129;
Matches  526;  Conservative 185;  Mismatches 408;  Indels 30;  Gaps 11;

Qy   13  KKILVANRGEIAVRAFRAALETGAANTVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL   72
         ||:|||||||||:|. || || :|  ||||||: :||||||: . :|::.| .| | |||||L
Db    5  KKVLVANRGEIAIRVMRACTELKIKTVAIYSQEDTGSFHRYKSDEAYLVGAGKKPIDAYL   64

Qy   73  DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT  132
         ||:|||  ||. .|||::||||||||| : |||| :.|| |:||.  .|::: ||| :|
Db   65  DIENIIEIAKSGADAIHPGYGFLSENIEFARRCEQEGIIFVGPKSKHLDMFGDKIKAKE  124

Qy  133  AAKKAGLPVLAEST-PSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKLA  191
         : ::| :|::..| ..::.. .:.| | .. |:::|:|: ||||||||:| |  : ||.
Db  125  QALLADIPVIPGSNGPVAGIKEVEEFGEKNGYPLMIKASLGGGGRGMRVVESKEHVKESF  184

Qy  192  TEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVEI  251
         .:||.|| ||||:. ||||: |:||:|||||||||||.::||||:|||||||||||||:|
Db  185  ERASSEAKAAFGNDEVYVEKCVMNPKHIEVQILGDHGNIVHLFERDCSIQRRHQKVVEV  244

Qy  252  APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE  311
         ||.:.:.:||.:|||.:||:. ::: |:|||||||||.|..::.||| |||:|||||:||
Db  245  APCNAITSELRNRICDAAVKLMKNVDYINAGTVEFLV-EGDDFYFIEVNPRVQVEHTITE  303

Qy  312  EVTEVDLVKAQMRLAAGATLKELGLT--QDKIKTHGAALQCRITTEDPNNGFRPDTGTI  368
         :.| :|:|::|:  |.|.:|::.|.:  ::::|:||:|:| ||||||| ||| ||||.:
Db  304  MITGIDIVQSQLFIADGYALHDQLVAIPKQEDIHIHGSAIQSRITTEDPLNNFMPDTGRV  363

Qy  369  TAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSG  427
         |.|||.||.|||||  ..||||.:.:.|:||||||: : | || |.:::::|:.||::.|
Db  364  DTYRSTGGFGVRLDAGNGFQGGTVVTPFYDSLLVKLCTWGMTFEQATRKMRRNLIEFRIRG  423

Qy  428  VATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDLADVTVNKPHG  487
         | ||: || . :|: || |.  .||||.::||| : :   | : |::.:.:.|||.  |
Db  424  VKTNIPFLLNVVRHPDFASGNYNTSFIDTTPELFKFPHIRDRGTKTLRYIGNVTVNGFPG  483
```

```
Qy    12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY    71
         ||:||||||||||:|:||||  :|:  || |||||:  :  |   |  ||  || ||:|||
Db     5 FKRYLVANRGEIAIRIFRACHELGIRTVAIYSEEEKLALFRTKADESYLLGQNKGPVDAY    64

Qy    72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV   131
         |:||||| ||||  ||||:|||||||||:: :|| | ||| |||||:::: |||||:|||
Db    65 LNIDEIINLALKKGVDAIHPGYGFLSENSEFSRRCTEAGIEFIGPTGDMMDKLGDKINSK   124

Qy   132 TAAKKAGLPVL-AESTPSKNIDEIVKSADGQTYPIFVKAVAGGGGRGMRFVASPDELRKL   190
         ||||:|| ::| |  : |  |||  |:    | :||||||||||||||||:||| |:|:
Db   125 LAAKAAGVKTIPGVEKPIETEQQAIEFARTCGYPVMVKAAAGGGGRGMRIVEKEEDLIAA   184

Qy   191 ATEASREEAEAFGDGAVYVERAVINPQHIEVQLLGDHTGEVHLYERDCSLQRRHQKVVE    250
         |  |   |||||| ||  :||||:||:|||||:|||:  |:||||||||:|||||||:|
Db   185 CRSAKSEAKKAFGLEDIFEKYLEGPKHIEVQVLGDKYGNIVHLYERDCSVQRRHQKVIE    244

Qy   251 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT   310
         :  | : || :|:|||:|||| :||||| :|||:||| |::|||:|||||||:|||||:|
Db   245 LTPAVSMSEEKRLEICEDALKIARSIGYRSAGTLEFLLDKHGNHYFIEMNPRVQVEHTIT   304

Qy   311 EEVTEVDLVKAQMRLAAGATLK-ELGL-TQDKIKTHGAALQCRITTEDPNNGFRPDTGT    367
         | || :|:|::|:|:|||: |  | |  :|:  |:::|::|||||||||:|:|  ||||
Db   305 EMVTGIDIVQSQILIAEGYKLNSPEVGINSQEDIHVNGYAIQCRITTEDPSNSFAPDTGK   364

Qy   368 ITAYRSPGGAGVRLDGA-AQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVS   426
         | | ::  | |:||||  ::  | :| :|||:||| |  | :||: |:||:|:|:|:|:|
Db   365 IDYVRTGSGFGIRLDGGNGFTGAVISPYYDSLLVKSTSWSRTFEDAIRKAIRAIKETYIS   424

Qy   427 GVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVNKPH   486
         || ||| ||  |:|| ||  :  |:  ||:|  :   |  || ::|: :  |    |:
Db   425 GVKTNIDFLINVLNHETFRKGLCDTNFIANNPLFEITPRIDTELRVLKFIGEKVVNETH    484

Qy   487 GVRPKDVAAPIDKLP--NIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDA   544
         |  :  : :| :|    :|  :| || :   :  :| | :: :|:::|  |:||||:||:
Db   485 GHK---IEVDVPSVPKYEIKE-PL-RGTKQILDEKGPKGLVEWIKDQDKLLLTDTTMRDA   539

Qy   545 HQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREA   604
         ||||:||| |: :::  |:|| ||  |  | | |||||:|:| || ||:||:||: |::|
Db   540 HQSLMATRLRTVDMVKIAKAESVLAKDLFSMEMWGGATFDTAYRFLKESPWERLERLKKR   599

Qy   605 MPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAV   664
         :|||::|||||| |:||| ||||:|  |||:: : ||:|:||:||: | |:::   |:|:
Db   600 VPNVLFQMLLRGANAVGYKNYPDNVIREFIKQSSKSGIDVFRIFDSLNWVKGMEVAIDEV   659

Qy   665 LETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAAV   724
         |  |  | | | : :||:  ||:|||||:|| :::|||:|:|||||| |||::||:| || 
Db   660 LNQG-KVAEAACMCYTGDILDTNRDKYTLNYYVNLAKEIEKSGAHILGKDMSALLKPYAA   718

Qy   725 TKLVTALRREFDLPVHVTHDIAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSAI   784
         :||: |||: :  |:|:|||: | |:||| : || ||:||| :| :: |:|||| ||:|
Db   719 LKLIRALKNEISIPIHLHTHDTTGNGVATVLMAAHAGVDIADTAFNSMSGLITSQPALNSV   778

Qy   785 VAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQLSNLR   844
         ||| ::| |||  ::  : |:  ||:||: :|  ||||   : |:| :|||||| |||:
Db   779 VAALKNTDRDTKMDIGDLQKISDYWSTVRPVYSKFESGLKAVSAEIYKYEIPGGQYSNLK   838

Qy   845 AQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAADP   904
         :| : | |  || ::::|:  || :|| ::|||||| |||||:::|| : ||:::  |:|
Db   839 PQVESFGLGHRFEQVKEMYREVNIMLGDIVKVTPSSKMVGDLAIFMVQNELTSENILEKA   898

Qy   905 QKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDADDSK   964
         ::  |||||: : ||::| |  | |:| :::  :| | |:  |: : :  :: | |:|:
Db   899 KDMPFPDSVVSYFKGMMGQPKGGFPKELQKIVL---KDEEAITCRPGELLPDEDFDKIR   954

Qy   965 ERRNSLNRL------LPFKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIR      1013
         :  :  ::|      |  : :: :|: | ::|: |  | :: :: || || :||
Db   955 VRLKKENKIEPTDKDVISYALYPDVFEEYLKYKNEYGDLSRMGSDVFFHGLAEGEISLE    1014

Qy  1014 LPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEK---ADSS   1070
         :  | | |||:|  |  :::|: | |: ||||:| |::: : ::|: ::   |   ||||
Db  1015 IAEGKT-LVVQLLHIGKLDKQGNRTLVFEVNGNRREIKIKDKVSSTKSEIVEIVIADSS    1073
```

```
Qy   488 VRPKDVAAPIDKLPNIKDLP----LPRGSRDRLKQLGPAAFARDLREQDA.LAVTDTTFRD   543
         :  |  : :| :|    |:         :| || :   :  :| | :: :|  |:||||:||
Db   484 IKHRD--KPVYAEPRLPKIPYGSQISPGTKQIILDAKGPEGVVDWVKKQEEVLLTDTTLRD   541

Qy   544 AHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELRE    603
         ||||:||| |: :::  |:|| ||  |  | | |||||:|:| || ||:||:||: |::
Db   542 AHQSLLATRVRSKDIFQVADAMAHLLPNMFSFEMWGGATFDVAYRFLNEDPWVRLETLRK    601

Qy   604 AMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDA   663
         :|||::|||||| |:||| ||||:|  |||:: : ||:|:||:||: | |:::   |:|
Db   602 QIPNVMFQMLLRGANAVGYKNYPDNVIREFVKQSAQSGVDVFRVFDSLNWIKGMEVSIDA   661

Qy   664 VLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAA   723
         |:  |  | | | : :||:  ||:|||||:|| :::|||:|:|||||| |||::||:| |
Db   662 VREAG-KVVEAAICYIGDIDDIRTKYTIDYYKDMAKELVAQGTHILGIKDMAGLLKPQA    720

Qy   724 VTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSA   783
         :||: |||: :  |:|:|||: | |:||| : || ||:||| :| :: |:|||| ||:
Db   721 AYRLIGELKDTVDVPIHLHTHDTSGNGIYTYAAAVSAGVDIVDVASSAMSGATSQPSMTG   780

Qy   784 IVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPESGTPGTGRVYRHEIPGGQLSNL    843
         |||| ::| |||  ::  : |:  ||:||: :|  ||||   : |:| :|||||| |||
Db   781 LYYGLVNGNRQTNLDAQNSQIINHYWNEDVRHYYKDFDNALNSPQTEVYIHEMPGGQYTNL  840

Qy   844 RAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAAD   903
         :| : | |  || ::::|:  || :|| ::|||||| |||||:::|| : ||:::  |:
Db   841 QQQAIAVGLGDRWDEVKEMYTVVNQMFGDIYKVTPSSKVVGDLALFMVQNESEEDVYEK    900

Qy   904 PQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDADDS   963
         |::  |||||: : ||::| |  | |:| :::  :| | |:  |: :  :  :: | |:|
Db   901 GDTIDFPDSVIEFFMGEIGQPYGGFPEPLQKLQKIVLKGRT---PLTDRPGALMEPVNFVDV 956

Qy   964 K------ERRNSLNRLLFFKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLI      1012
         |      :  :  ::| |  : :: :|: | ::|: |  | :: :: || || :|
Db   957 KAELKEKMGYEPTEKDVISYILYPKVFLDYQDMINKYGDVTVLDTPTFYKGMRLGETIEV   1016

Qy  1013 RLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNK   1072
         ::  | | |||:|  |  :::|: | |: ||||:| |::: : ::|: ::   |: :::|
Db  1017 ELEKGKI-LLIKLNSIGEPIADGTRVIYFELNGQPREINIQDMNVQSTVIARRKIDTTNP   1075

Qy  1073 GHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATTTASVDGKIDRVVVPAATKVE   1131
         ||| | : | |:|:|| :|| ||||| :|||||||| :|| | :  |:|:|: | :| :
Db  1076 EHVGATMTGSVIQVVVKKGDSVKKGDPLLITEAMKMETTIQAPFDGEVSSIIVSDGDTIE   1135

Qy  1132 GGDLIVVVS   1140
         ||||:: :
Db  1136 SGDLLIEVN   1144
```

RESULT 8
D97227
Pyruvate carboxylase, PYKA [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C;Accession: D97227
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: D97227
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1144 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK80607.1; PID:g15025689; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC2660
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

Query Match   43.0%; Score 2491.5; DB 2; Length 1144;
Best Local Similarity 45.3%; Pred. No. 5e-129;
Matches 520; Conservative 204; Mismatches 391; Indels 33; Gaps 14;

```
Qy  1071 NKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVPAATK 1129
         ||| ::|| ::| |||| : | ||| ::||||||:  || |||||||||||||||:
Db  1074 NKKEIGASIPGNVVKFVFKPGDKVKKGDSLMVIEAMKMETNVSVSEDGTVGGIFVKEGDQ 1133

Qy  1130 VEGGDLIV 1137
         | :|:|:|
Db  1134 VQSGQLLV 1141

RESULT 9
AC1565
Pyruvate carboxylase homolog pycA [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AC1565
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.;
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss,. H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:2157279; PMID:11679669
A:Accession: AC1565
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1146 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC96291.1; PID:g16413519; GSPDB:GN00178
C:Genetics:
A:Gene: pycA
C:Superfamily: pyruvate carboxylase; biotin carboxylase homology: lipoyl/biotin-binding

Query Match      43.0%;  Score 2490;  DB 2;  Length 1146;
Best Local Similarity  46.0%;  Pred. No. 6.1e-129;
Matches  528;  Conservative 182;  Mismatches 409;  Indels 30;  Gaps 12;

Qy    13 KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL 72
         ||:||||||||:|..||| :||::|:||||||:|: :.|||| ||| |:::|:: |:|||
Db     5 KKVLVANRGEIAIRVMRACTELKIKTVAIYSQEDTGSFHRYKSDEAYLVGAGKKPIDAYL 64

Qy    73 DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT 132
         ||:|:|  ||: |||||:|||||||||.: ||.|.:.|| |:||: :.|| :|||. ||.
Db    65 DIENIIEIAKESGADAIHPGYGFLSENIEFARRCEQEGIIFVGPKSKHLDMFGDKIKAKE 124

Qy   133 AAKKAGLPVIAEST-PSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRFVASPDELRKLA 191
         .| ||:.||: |:  . :|::|:.. |.|| :|::||| ||||:||||.|:. :: |:.:
Db   125 QALLADIPVIPGSDGVAGIKEVEFGEKNGYPLMIKASLGGGGRGMRVVESKEHVKESF 184

Qy   192 TEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVEI 251
         .:||.||:|||.: .||||..|:||:||||||||| .|:.|||.|||||:||||||||:
Db   185 ERASSEAKAAFGNDEVYVEKCVMNFKHIEVQILGDTHGNTVHLFERDCSIQRRHQKVVEV 244

Qy   252 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE 311
         ||.: : .||.:||.|||||:.. :.|::.|||||:| ||:.::|||:||:||||||:||
Db   245 APCNAITSELRNRICDAAVKLMKNVDYINAGTVEFLV-EGDNFYFIEVNPRVQVEHTITE 303

Qy   312 EVTEVDLVKAQMRLAAGATLKELGLT---QDKIKTHGAALQCRITTEDPNNGFRPDTGTI 368
         :.| :|:|::|:|:|.|.:| :|.|    :: ::||.::||.||||||| |.| |||| :
Db   304 MITGIDIVQSQLFIADGYALHDQLVAIPKQEDIHIHGSAIQSRITTEDPLNNFMPDTGRV 363

Qy   369 TAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSG 427
         .:||| ||.||||  |..: .| :|.|:||:|||:.| |. ||:|| :.:|.|||| . |
Db   364 DTYRSTGGFGVRLDAGNGFQGTVVTPFYDSLLVKLCTWGMTFEQATRKMRRNLIEFRIRG 423

Qy   428 VATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTNKPHG 487
         |.||:.||  :.|..||.|..: |.||||:. .|. | |..: . :|: .  ||.:. |
Db   424 VKTNIPFLLNVVRHPDFASGNYNTSFIDTTPELFKFPHIRDRGTKTLRYIGNVTVNGFPG 483

Qy   488 VRPKDVAAPIDKLPNIKDLP---LPRGSRDRLKQLGPAAFARDILREQDALAVIDTTFRD 543
         :: :|: | |.:||:|  |    :| |:. ::::|| . :..::. ::::|||||||:||
Db   484 IKHRD--KPVYAEPRLPKIPYGSQIAPGTKQILDAKGPEGVVDWVKKQEEVLLIDTTLRD 541
```

```
Qy   544 AHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLPEDPWDRLDELRE 603
         ||||||||||||  :. ||:.:|.|.|:::. |:||||| |||:||| |||| ||: ||:
Db   542 AHQSLLATRVRSKDIFQIADAMAHLLPNMFSFEMWGGATFDVAYRFLNEDPWVRLETLRK 601

Qy   604 AMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDA 663
          :|||.:||||||.|:||| .|||.| |.|||::| |||||||:||:||.: . ||:|||
Db   602 QIPNVMFQMLLRGANAVGYKNYPDNVIREFVKQSAQSGVDVFRVFDSLNWIKGMEVSIDA 661

Qy   664 VLETNTAVEVAMAYSGDLSDPNEKLYTLDYLYKMAEEIVKSGAHILATKDMAGLLRPAA 723
         | |   .||.|. ..||:.|  .   |:|| .|| ::||::|:|||. ||||||| ||
Db   662 VREAG--KVVEATICYTGDIDDTRTKYTIDYKDMAKELVAQGTHILGIKDMAGLLKPQA 720

Qy   724 VTKLVTALRREFDLPVHVTHDTAGGQLATYPAAAQAGADAVDGASAPLSGTTSQFSLSA 783
         . :|: :|:::.|:|:|:.||||:|   |.|||..||.|:||.|| |:||||||.|::.
Db   721 AYRLIGELKDTVDVPIHLHTHDTSGNGIYTAAAVSAGVDIVDVASSAMSGATSQPSMTG 780

Qy   784 IVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFRSGTPGTGRVYRHEIPGGQLSNL 843
         : . . . .:.  :.:.::.: :.:::.||.::   :...   .:.::.::  ||:.||
Db   781 LYYGLVNGNRQTNLDAQNSQIINHYWEDVRHYYKDFDNALNSPQTEVVIHEMPGGQYTNL 840

Qy   844 RAQATALGLADRFELIEDYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAAD 903
         :.|| |.|| ||::::::.| ||:|.|.:.:||||||||||||| :  . :: .:   .
Db   841 QQQAIAVGLGDRWDEVKEMYTVVNQMFGDIVKVTPSSKVVGDLALFMWQNELTEEDVYEK 900

Qy   904 PQKYDIPDSVIAFLRGELGNPPGGWPRPLRTRALEGRSEGKAPLTEVP---------EEE 954
          : .||||||| ::.||:|:| ||:|. |:  .|:|:  ...||.: |         .
Db   901 GDTIDFPDSVIEFFMGEIGQPYGGFPEKLQKLVLKGRT----PLADRPGALMEPVNFAEV 956

Qy   955 QAHLDADDSKE--RRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLI 1012
          | | :.. .:  . .|:. :::.||..::::     ||:: .|. .  ::|: .|||:
Db   957 KAELKEKMGYEPSEKDVISYILYPKVFLDYQEMISKYGDVTVLDTPTFYKGMRLGETIEV 1016

Qy  1013 RLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATEKADSSNK 1072
          | : : .||:.|::|.||.|:|.|.: .:::|| | ::: : .: . | .. | :.|
Db  1017 ELEKGKI-LLIKLNSIGEPIADGTRVIYFELNGQPREINIQDMNVQSTVIARRKIDTTNP 1075

Qy  1073 GHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVPAATKVE 1131
          ||.|.: | ::.|.|.:|| |||||.: |||||||||:|  | |||:. .::.|| .:
Db  1076 EHVGAIMTGSVIQVVVKKGDSVKKGDPLLITEAMKMETTIQAPFDGEVSSIYVSDGDTIE 1135

Qy  1132 GGDLIVVS 1140
          ||| ::|
Db  1136 SGDLLIEVN 1144

RESULT 10
G89881
Pyruvate carboxylase [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: G89881
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: G89881
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1150 <KUR>
A:Cross-references: GB:BA000018; PIDN:BAB42211.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: pycA
C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindi

Query Match      43.0%;  Score 2488;  DB 2;  Length 1150;
Best Local Similarity  45.2%;  Pred. No. 7.9e-129;
Matches  527;  Conservative 192;  Mismatches 384;  Indels 62;  Gaps 14;

Qy    13 KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL 72
         ||:||||||||| |||| ||:||.|:|:|| |  |: :|:|| ::  |:.|::| |:|:|
```

```
Db      5  KKLLVANRGEIAIRIFRAAAELDISTVAIYSNEDKSSLHRYKADESYLVGSDLGPAESYL   64
           :| ||::|:| |:|:||||::  ::  ::|: :||  |||: :|||||||: |: | : |
Qy   1115  VDGKIDRVVPAATKVEGGDLIVVV  1139

Qy     73  DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT   132
           :|:|| ||||  ||||||||||||||||   || |||| ||: :|||
Db   1121  FDGVIKQVTVNNGDTIATGDLLIEI   1145

Db     65  NIERIIDVAKQANVDAIHPGYGFLSENEQFARRCAEEGIKFIGPHLEHLDMFGDKVKART  124

RESULT 11
AE2911
pyruvate carboxylase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AE2911
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1174 <KUR>
A:Cross-references: GB:AE008688; PIDN:AAL43707.1; PID:g17741236; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: pycA
A:Map position: circular chromosome
C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindi

Query Match             42.7%;    Score 2473.5;   DB 2;     Length 1174;
Best Local Similarity   46.6%;    Pred. No. 5.1e-128;
Matches  541;  Conservative 173;  Mismatches 406;   Indels  41;   Gaps  19;


Qy      8  TLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIG------   62
           | |||||:||||||||| |||| |||||| ::|| ::: : |:  |::
Db     20  TVLKISKILVANRSEIAIRVFRAANELGIKTVAIWAEEDKLSLHRFKADESYQVGRGPHL   79

Qy     63  -TEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVL  121
            : |||| : |||: ||:|  ||||| | |||||||||||:: |||  | | ||||||:|
Db     80  AKDMGPIESYLSEEVIRVAKLSGADAIHPGYGLLSESEPFVEACNKAGITFIGPTPDTM  139

Qy    122  DLTGDKSRAVTAAKKAGLPVLAESTP-SKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF  180
           |||| | |||| || |  || ||  | |:  :| |:   | || ::||||| |||||:
Db    140  ROLGNKVAARNLAISVDVPVVPATNPLPDDIAEVERMAEEIGYPVMLKASWGGGGRGMRA  199

Qy    181  VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS  240
           || | |:|  || |||  || ||   |||| | |:| :|: |:  |||  |||||||||:
Db    200  IRKKEDILAREVTEAKREAKAAFGKDEVYLEKLVERARHVESQILGDTHGNVVHLFERDCS  259

Qy    241  LQRRHQKVVETAPAQHLDPELERDRICADAVKPCRSIGYQGAGTVEFLVD-EKGNHVFIEM  299
           :|||:|||: | || :| | ||:| |:   : ||::|  || ||| || ||| :| : :
Db    260  IQRRNQKVVERAPAYLSEAQRQELAAYSLKIAAATNYIGAGTVEYLMDADFGKFYFIEV  319

Qy    300  NPRIQVEHTVTEEVTEVDLVKAQMRLAAGATL--KELGL-TQDKITHGAALQCRITTED  356
           ||||||||||||||||||||||||||||||||  :|||| ||||||||||:|||||||
Db    320  NPRIQVEHTVTEVTGIDIVKAQIHLLEGAAIGTAESGVPKQEDIRLNGHALQCRITTED  379

Qy    357  PNNGFRPDTGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVAR  415
           | : |  : :| |||: || |:|||  | | || |  :| : :|  ||: | :: :|:
Db    380  PEHNFIPDYGRITAYRSASGFGIRLDGGTSYTGAVITRYYDPLLVKVTAWAPEPDEAISR  439

Qy    416  AQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILD  475
           | | | :: : :||:  |::  | :|::: :: ||:   :   ::|| :|   ::: :
Db    440  MDRALREFRIRGVATNLTFLEAIIGHDSFRNNTYTTRFIDSTPELFAQVKRQDRATKLLT  499

Qy    476  YLADVTVN--KPHG-VRPKDVAA-PIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLRE  530
           |||| | |  |||  ||    | | :| ::|  |  | |   :   | ||  |   | :
Db    500  YLADVTVNGHPETKGRAKPADKAAKPI--VPYI-DAPTPDGTKQLLDKLGPKGFADWMRN  556

Qy    531  QDALAVDTTFDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFL  590
           : : ||||||||||||||||||||||||||| |  | :||||:|||||||  :||:|:
```

```
Db     73  DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT  132
Qy   1121  FDGVIKQVTVNNGDTIATGDLLIEI  1145
Db     65  NIERIIDVAKQANVDAIHPGYGFLSENEQFARRCAEEGIKFIGPHLEHLDMFGDKVKART  124
Qy    133  AAKKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKLA  191
Db    125  TAIKADLPVIPGTDGPIKSYELAKEFAEEAGFPLMKATSGGGGKGMRIVREESELEDAF  184
Qy    192  TEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVEI  251
Db    185  HRAKSEAEKSFGNSEVYIERYIDNPKHIEVQVIGDEHGNIVHLFERDCSVQRRHQKVVEV  244
Qy    252  APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLV--DEKGNHVFIEMNPRIQVEHTV  309
Db    245  APSVGLSPTILRQRICDDAAIQLMENIKYVNAGTVEFIVSGDE--FFFIEVNPRVQVEHTI  301
Qy    310  TEEVTEVDLVKAQMRLAAGATL--KELGLTQDK-IKTHGAALQCRITTEDPNNGFRPDTG  366
Db    302  TEMVTGIDIVKTQIIVAAGADLFGEEINMPQQKDITTLGYAIQCRITTEDPLNDEMPDTG  361
Qy    367  TITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTV  425
Db    362  TIIAYRSSGGFGVRLDAGDGFQGAEISPYYDSLLVKLSTHAISFKQAEEKMVRSLREMRI  421
Qy    426  SGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVNKP  485
Db    422  RGVKTNIPFLINVMKNKFTSGDYTTKFIEETPELFDIQPSLDRGTKTLEYIGNVTIN--  479
Qy    486  HGVRPKDVAAPIDKLPNIKDLPLP-----------RGSRDRLKQLGPAAFARD  527
Db    480  ---------GFPNVEKRPKPDYELASIPTVSSSKIASFSGTKQLLDEVGPKGVAEW  526
Qy    528  LREQDALAVDTTFDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAM  587
Db    527  VKKQDDVLLTDTTFRDAHQSLLATRVRTKDMINIASKTADVFKDGFSLEMWGGATFDVAY  586
Qy    588  RFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRI  647
Db    587  NFLKENPWERLERLRKAIPNVLFQMLLRASNAVGYKNYPDNVIHKFVQESAKAGIDVFRI  646
Qy    648  FDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDP-NEKLYTLDYYLKMAEEIVKSG  706
Db    647  FDSLNWVDQMKVANEAVQEAG-KISEGTICYTGDILNPERSNIYLEYYLYLAKELEREG  705
Qy    707  AHILAIKDMAGLLRPAAVTKIVTALRREFPLPVHVTHDTRAGGQLATYFAAAQAGADAVD  766
Db    706  FHILAIKDMAGLLKPKAAYELIGELKSAVDLPIHLHTHDTSGNGLLTYKQAIDAGVDIID  765
Qy    767  GASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGP  826
Db    766  TAVASMSGLTSQPSANSLYYALNGPRHLRTDIEGMESLLSHYWSTVRTYYSDFESDIKSP  825
Qy    827  TGRVYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDL  886
Db    826  NTEIYQHEMPGGQYSNLSQQAKSLGLGERFDEVKDMYRRVNFLFGDIVKVTPSSKVVGDM  885
Qy    887  ALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPGGWPEPLRTRALEGRSEGKAP  946
Db    886  ALYMVQNDLDEQSVITDGYKLDFPESVVSFFKGEIGQPVNGFNKDLQAVLLKGQE----A  941
Qy    947  LTEVPEEEQAHLDADDSKE---------RRNSLNRLLFPKPTEEFLEHRRRFGNTSAL  995
Db    942  LTARPGEYLEPVDFEKVRELLEEEQQGPVTEQDIISYLLPKVVEQYIQTRNQYGNLSLL  1001
Qy    996  DDREEFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDR  1055
Db   1002  DTPTFFGMRNGETVEIEI-DKGKRLIIKLETISEPDENGNRTIYYAMNGQARRIYIKDE  1060
Qy   1056  SVESVTATAEKADSSNKGHVAAPFAGVVT-VTVAEGDEVKAGDAVAIIEAMKMEFATITAS  1114
Db   1061  NVHTNANVKPKADKSNPSHIGAQMPGSVTEVKVSVGETVKANQPLLITEAMKMETTIQAP  1120
```

```
                    :::||| ||||: ||||   ||| ||||  ||:|
Db   80 AKDMGPIESYLSIEEVIRVAKLSGADAIRPGYGLLSESPEFVEACNKAGTIFIGPTPDTM 139
Qy  122 DLTGDKSRAVTAAKKAGLPVLAESTP-SKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRF 180
Db  140 RQLGNKVAARNLAISVDPVPVPATNPLPDDIAEVERMAEEIGYPVMLKASWGGGRGMRA 199
Qy  181 VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVHLYERDCS 240
Db  200 IRKKEDLAREVTEAKREAKAAFGKDEVYLEKLVERARHVESQILGDTHGNVHLFERDCS 259
Qy  241 LQRRHQKVETAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVD-EKGNHVFIEM 299
Db  260 IQRRNQKVERAPAPYLSEAQRQELAAYSLKIAAATNYIGAGTVEYLMDADTGKFYFIEV 319
Qy  300 NPRIQVEHVTEEVTEVDLVKAQMRLAAGATL-KELGL-TQDKIKTHGAALQCRITTED 356
Db  320 NPRIQVEHTVTEVVTGIDIVKAQIHILEGAIGTAESGVPKQEDIRLNGHALQCRITTED 379
Qy  357 PNNGFRPDTGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVAR 415
Db  380 PEHNFIPDYGIRITAYRSASGFGIRLDGGTSYTGAVITRRYDYDPLLVKVTAWAPEPDEAISR 439
Qy  416 AQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILD 475
Db  440 MDRALREFRIRGVATNLTFLEAIIGHDSFRNNTYTTRFIDSTPELFAQVKRQDRATKLLT 499
Qy  476 YLADVTVN---KPHG-VRPKDVAA-PIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLRE 530
Db  500 YLADVTVNGHPETKGRAKPADKAAKPI--VPYI-DAPTPDGTKQLLDKLGSPKGFADWMRN 556
Qy  531 QDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFL 590
Db  557 EKRVLVTDTTMRDGHQSLLATRVRTHDIARVASVYSKALPQLLSLECWGGATFDVSMRFL 616
Qy  591 FEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIRIFDA 650
Db  617 TEDPWERLSLIREGAPNLLLQMLLRGANGVGYKNYPDNVVKYFVRQAARGGVDLFRVFDC 676
Qy  651 LNDVSQMRPAIDAVLETNIAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHIL 710
Db  677 LNWVENMRVSMDAIAEEN-KLCEATICYTGDLLNSARPKYDLKYYTNLAVELEKAGAHII 735
Qy  711 AIKDMAGLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASA 770
Db  736 AVKDMAGLLKPAAAKVLFKALREATGLPIHFHTHDTSGISAATVLAAVDAGVDAVDAAMD 795
Qy  771 PLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFEESGTPGTGRV 830
Db  796 AFSGNTSQPCLGSIVEALSGSERDTGLDTEWIRRISFYWEAVRNQYAAFESDLKGPASEV 855
Qy  831 YRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHL 890
Db  856 YLHEMPGGQFTNLKEQARSLGLESRWHEVAQAYADANRMFGDIVKVTPSSKVVGDMALMM 915
Qy  891 VGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEV 950
Db  916 VSQDLTVADVENPDREVSFPDSVVSMLKGDLGQSPGGWPEALQKKAL---KGEKPYTVR 971
Qy  951 PEE--EQAHLDADDSKERRNSLNR-------LLFPKPTEEFLEHRRRFGNTSALDDR 998
Db  972 PGSLLEDADLDA-ERKVIETKLERKVDDFEFASYLMYPKVFTDFALTAETYGPVSVLPTH 1030
Qy  999 EFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRS-V 1057
Db 1031 AYFYGMEDGEELFADIERGKTLVIVN-QASSGIDDKGMVTVFFEINGQPRRIKVPDRAHG 1089
Qy 1058 ESVTATAEKADSSNKGHVAAPFAGVVT-VTVAEGDEVKAGDAVAIIEAMKMEATITASVD 1116
Db 1090 ASGSAVRRKAEPGNASHIGAPMPGVISRVFINQGQEVKAGDVLLSIEAMKMETALHAERD 1149
Qy 1117 GKIDRVVVPAATKVEGGDLIV 1137
Db 1150 GKIAEVLVKPGDQIDAKDLLI 1170
```

RESULT 12
C97686
Pyruvate carboxylase (U51439) [imported] - Agrobacterium tumefaciens (strain C58, Cereon…
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C;Accession: C97686
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
  A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: C97686
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1174 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK88444.1; PID:g15157941; GSPDB:GN00169
C;Genetics:
A;Gene: AGR_C_4940
A;Map position: circular chromosome
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology: lipoyl/biotin-binding

Query Match   42.7%;   Score 2473.5;   DB 2;   Length 1174;
Best Local Similarity   46.6%;
Matches 541;   Conservative 173;   Mismatches 406;   Indels 41;   Gaps 19;

Pred. No. 5.1e-128;

```
Qy    8 TLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIG------ 62
Db   20 TVLIKISKILVANRSEIAIRVFRAANELGIKTVAIWAEEDKLSLHRFKADESYQVGRGPHL 79
Qy   63 -TEGSPVKAYLDIEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTEVL 121
```

```
Db   1150 GKIAEVLVKPGDQIDAKDLLI 1170

RESULT 13
AE3285
Pyruvate carboxylase (EC 6.4.1.1) [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C;Accession: AE3285
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
  Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
  Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AE3285
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1158 <KUR>
A;Cross-references: GB:AE008917; PIDN:AAL51448.1; PID:g17982157; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI0266
A;Map position: I
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding
C;Keywords: ligase

Query Match         42.6%;  Score 2464.5;  DB 2;  Length 1158;
Best Local Similarity 46.6%;  Pred. No. 1.6e-127;
Matches  538;  Conservative 172;  Mismatches 408;  Indels 37;  Gaps 15;

Qy   14 KILVANRGEIAVRAFRAALETGAAIVAIYPREDRGSFHRSFASEAVRIG-----TEGSP 67
         ||||||| ||| | ||||| || | | |||   | : | | : :  :         :|::
Db   10 KILVANRSEIAIRVFRAANELGLKTVTIWAEEDKLSLHRFKADESYQVGRGPHLDRDLGP 69

Qy   68 VKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPPEVLDLTGDK 127
         :  :||||||  |  : ||||| ||||| || |  | | |||| | ||| | : :| :|
Db   70 IESYLSIDEIIRVAKLSGADAIHPGYGLLSESPEFAEACAENGIVFIGPKPETMRRLGNK 129

Qy  128 SRAVTAAKKAGLPVLAESTP-SKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRFVASPDE 186
         .  . || || ||: . :   . :|||| : | :|| |:::||: |||||||| : :::
Db  130 VAARNLAIEIGVPVVPATDPLPDDMDEVKKLAAQIGYPLMLKASWGGGGRGMRAIRAEAD 189

Qy  187 LRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQ 246
         | ::  ||   | ||||  | |:| ||| ::||||||||  ||   ||:||||| |||.|
Db  190 LAREVWEAKRAKAAFGKDEVYLEKLVERARHVEVQILGDTYGNAVHLFERDCSIQRRNQ 249

Qy  247 KVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVD-EKGNHVFIEMNPRIQV 305
         ||||   | | | : | |  |: | : ||| | |||||||:.|   |: ||||:||||||
Db  250 KVVERAPAPYLNDAQRRELADYGLKIAHATDYIGAGTVEFLMDADTGKFYFIEVNPRIQV 309

Qy  306 EHTVFBEVTEVDLVKAQMRLAAGATL--KELGL-TQDKIKTHGAALQCRITTEDPNGFR 362
         |||| ||| : |:|||: |  | |:   |: |: . |:|  |||||||||||||| : |
Db  310 EHTVFEEVTGIDIVKAQIHILEGFAIGTPESGVPRQEDIRLNGHALQCRITTEDPEQNFI 369

Qy  363 PDTGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALA 421
         || | | ||||. | |:|||    : .||| |:|||:|  ::  |:   | | |:  ||
Db  370 PDYGRIQAYRSAAGFGIRLDGGTAYSGAFITRYYDSPLLVVKVTASGATPLEAIHRMDRALR 429

Qy  422 EFTVSGVATNIGFLRALLRREDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVT 481
         ||   |||| ||  |:   : : ||  | :|: ||:|: |:  :||: |  |  | |:||
Db  430 EFRIRGVATNLTFLEAIINHPKFLSNDYTTRFIDTTPELFEQMKRQDRATKLLIYIADVT 489

Qy  482 VN-----KPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAV 536
         ||      |  | . | :  .|:  :|  :  | |::  :  ||| :|| : |: |  |
Db  490 VNGHPETKGRAKPARDAAKP--RVPWFGDKLVADGTKQLLDQLGPKKFAEWMRNEKRALI 547

Qy  537 TDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWD 596
         |||| || ||||||||||: ::   | |: | ||:| |  ||| ||:|| |||  |||:
Db  548 TDTTMRDGHQSLLATRVRTYDIARIANAYAQALPNLFSLECWGGATFDVSMRFLTEDPWE 607

Qy  597 RLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQ 656
         || : ||| ||: :||||||   ||| .|||.| : ||||| :|:| ||:||:||::. :
Db  608 RLALVREGAPNLLLQMLLRGANGVGYKSYPDNVVKYFVREAARAGIDLFRVFDSLNWVEN 667

Qy  657 MRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLKMAEEIVKSGAHILAIKDMA 716
         ||  :||||| | ||:|  | | || |:|: :  |||| :: |:||:|||||:|:||||
Db  668 MRVSMDAVLEEN-KICEAAICYTGDILNPDRAKTDLNYYVNLAKEVEKAGAHIIAVKDMA 726

Qy  717 GLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTT 776
         ||| ||| : |  || || |||:| ||||| | : |: ||| ||| ::|: | | ||:|
Db  727 GLLKPAAARVLFKALREETDLPIHFHTHDTSGISAATVLAAIDAGVVDVDAAMDALSGNT 786

Qy  777 SQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIP 836
         |||.| :||:  ||:||: ||.  :   ||:|:|||| | : ||   | ::| | ||:|
Db  787 SQPCLGSIVEALHGSERDSGLDPDLIRRISFYWEAVRHQYAAFESDLKGPASEVYLHEMP 846

Qy  837 GGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVD 896
         ||| :||: || :|||  |: :: : |: ||:| | :||||||||||||:||::| |  |
Db  847 GGQFTNLKEQARSLGLETRWHEVAQAYADVNRMFGDIVKVTPSSKVVGDMALMMVAQDLT 906

Qy  897 PADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEE--E 954
          || ||| | | |||||:: |||:|| | ||||| |:  ||:| | : || || |:   |
Db  907 VADVENPAKDIAFPDSVVSMRGDLGQPPSGWPEALQKKVL----KDEKPFTVRPGSLLP 962

Qy  955 QAHLDA-------DDSKERRNS---LNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLV 1005
          | |||       :||  :  .    :|:|:||| |:|| || | :|:|:   : || |:|
Db  963 AADLDAERKSFEDSVGRKLSDQEFASALMYPKYFTDYATAHETYGPTSVLPTPVYFYGLK 1022

Qy 1006 EGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAE 1065
          :   .: : |: :|||:|  : :|||:||| .||  :||| |:::| :: : .  . .
Db 1023 PEEEVFVDLERGKTLVIVN-QAMSETDEKGMVTVFFELNGQPRRIKVPNRAKGASGGVRR 1081

Qy 1066 KADSSNKGHVAAPFAGVV-TVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVV 1124
          | .:::  | |||  || || || |: |  ||.| : ||||||: | :| : ||| :|:|
Db 1082 KVEAGNDKQVGAPMPGVISTVAVVAGQKVTQGDVLLSIEAMKMETAIHAERDGTIAEVLV 1141

Qy 1125 PAATKVEGGDLIVV 1138
          .  ::| ||:|:||
Db 1142 RPGEQIDAKDLLIV 1155

RESULT 14
T20346
pyruvate carboxylase (EC 6.4.1.1) D2023.2 [similarity] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000
C;Accession: T20346
R;Kershaw, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19260
A;Accession: T20346
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1175 <WIL>
A;Cross-references: EMBL:Z81052; PIDN:CAB02872.1; GSPDB:GN00023; CESP:D2023.2
A;Experimental source: clone D2023
C;Genetics:
A;Gene: CESP:D2023.2
A;Map position: 5
A;Introns: 34/3; 103/3; 246/1; 451/3; 567/2; 821/1; 1045/3; 1126/3
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindi
C;Keywords: ligase
F;1140/Binding site: biotin (Lys) (covalent) #status predicted

Query Match         42.6%;  Score 2464;  DB 2;  Length 1175;
Best Local Similarity 45.2%;  Pred. No. 1.7e-127;
Matches  523;  Conservative 200;  Mismatches 389;  Indels 46;  Gaps 14;

Qy   12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPYKAY 71
         | |::|||||||| | |||  |  || :|||  |: .| | |  :| |: ||  .| :||
Db   32 FNKVMVANRGEIAIRVFRALTELNKTSVAITAEQDKNSMHRLKADEAYLVGKGLPPVAAY 91

Qy   72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 131
         | ||:II : : : ||AI|PGYGFLSE :  | ||  |||.FIGP.P:V: : ||| | |
Db   92 LTIDQIIETALKHNIDAIHPGYGFLSERSDFAAACQNAGIVFIGPSPDVMARMGDKVAAR 151
```

RESULT 15

QYBYP

pyruvate carboxylase (EC 6.4.1.1) 1 [validated] - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein G3428; protein YGL062w; pyruvic carboxylase
C;Species: Saccharomyces cerevisiae
C;Date: 31-Dec-1991 #sequence_revision 19-Jul-1996 #text_change 01-Feb-2002
C;Accession: S64066; A29233; S05760; A29722
R;Feuermann, M.; Potier, S.; Souciet, J.L.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64044
A;Accession: S64066
A;Molecule type: DNA
A;Residues: 1-1178 <FEU>
A;Cross-references: EMBL:Z72584; NID:g1322565; PIDN:CAA96765.1; PIDN:g1322566; GSPDB:G
A;Experimental source: strain S288C
R;Lim, F.; Morris, C.P.; Occhiodoro, F.; Wallace, J.C.
J. Biol. Chem. 263, 11493-11497, 1988
A;Title: Sequence and domain structure of yeast pyruvate carboxylase.
A;Reference number: A92662; MUID:88298805; PMID:3042770
A;Accession: A29233
A;Molecule type: DNA
A;Residues: 1-461,'G',463-492,'D',494-594,'A',596-618,'Q',620-663,'S',665-771,'R',773
A;Cross-references: EMBL:J03889; NID:g172101; PIDN:AAA34843.1; PID:g172102
R;Morris, C.P.; Lim, F.; Wallace, J.C.
Biochem. Biophys. Res. Commun. 145, 390-396, 1987
A;Title: Yeast pyruvate carboxylase: gene isolation.
A;Reference number: S05760; MUID:87241529; PMID:3036126
A;Accession: S05760
A;Molecule type: DNA
A;Residues: 1003-1178 <MOR>
A;Cross-references: EMBL:J03889
A;Accession: A29722
A;Molecule type: protein
A;Residues: 1124-1149 <MOR2>
C;Genetics:
A;Gene: SGD:PYC1; PYV; MIPS:YGL062w
A;Cross-references: SGD:S0003030; MIPS:YGL062w
A;Map position: 7L
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindi
C;Keywords: biotin binding; gluconeogenesis; homotetramer; ligase; zinc
F;21-478/Domain: biotin carboxylase homology <BCH>
F;157-331/Domain: ATP/bicarbonate binding #status predicted <ATB1>
F;353-468/Domain: ATP/bicarbonate binding #status predicted <ATB2>
F;569-908/Domain: pyruvate binding #status predicted <PYR>
F;1096-1169/Domain: lipoyl/biotin-binding homology <LPB>
F;1135/Binding site: biotin (Lys) (covalent) #status experimental

Query Match   42.4%;   Score 2457;   DB 1;   Length 1178;
Best Local Similarity   46.7%;
Matches   541;   Conservative   173;   Mismatches   401;   Indels   44;   Gaps   16;
    Pred. No. 4.2e-127;

```
Qy     14 KILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFPHRSEAVRIGTEG--SPVKAY  71
          |||| :|| :||| ::||||| |::|| :|||||||| :| |:| ||  :|| ||| |
Db     21 KILVANRGEIPIRIFRTAHELSMQTVAIYSHEDRLSTHKQKADEAYVIGEVGQYTPVGAY  80

Qy     72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLLTGDKSRAV 131
          ||||||:||::|||||||||||||||||||::|: ||  :||||||||||  ||||:|
Db     81 LAIDEIISIAQKHQVDFIHPGYGFLSENSEFADKVVKAGITWIGPAEVIDSVGDKVSAR 140

Qy    132 TAAKKAGLPVLAEST-PSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRFVASPDELRKL 190
          |||||||||||||:| || |||:||||||||||| ||||||||||||||| |||||||:
Db    141 NLAAANVPTVPGTPGIEIVEEALDFVNEYGYPVIIKAAFGGGGRGMRVVREGDDVADA  200

Qy    191 ATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE 250
          |||||||| || |:|||| |||:||| |||||||||:||||:||||||| :||||||||:
Db    201 FQRATSEARTAFGNGTCFVERFLDRKPKHIEVQLLADNHGNVVHLFERDCSVQRRHQKVVE 260

Qy    251 IAPAQHLDPELRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 310
          |||||||||||||| ||| |:|||||||| ||||||||:||||||||||||||||||||
Db    261 VAPAKTLPREVRDALLTDAVKLAKECGYRNAGTAEFLVDNQNRHYFIEINPRIQVEHTIT 320

Qy    311 EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 370
          ||| :|| :|| || :|||||||||||:|||:|||||| ||:|||||| |||  | ||:|
```

```
Db    321  EEITGIDIVAAQIQIAAGASLPQLGLFQDKITTRGFAIQCRITTEDPAKNFQPDTGRIEV  380

Qy    371  YRSPGGAGVRLDGA-AQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVA  429
                || | |||||      :  ||||| :|  |:    ||  ||  |  :
Db    381  YRSAGGNGVRLDGGNAYAGTIISPHYDSMLVKCSCSGSTYEIVRRKMIRALIEFRIRGVK  440

Qy    430  TNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVN----KP  485
                || ||||  |     :     || | |||| :  : :|||||  || |||       ||
Db    441  TNIPFLLLTLLTNPVFIEGTYWTTFIDDTPQLFQMVSSQNRAQKLLHYLADVAVNGSSIKG  500

Qy    486  HGVRPKDVAAPIDKLPNIKDL---------PLPRGSRDRLKQLGPAAFARDLREQDALA  535
                ||  |  :  ||  |  |            |  |     ||  | |  : :|:  :
Db    501  QIGLPKLKSNP--SVPHLHDAQGNVINVTKSAPPSGWRQVLLEKGPAEFARQVRQFNGTL  558

Qy    536  VTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPW  595
                :  ||:|||||||||||||:   |    |    ::| |||||  ||||||||||||||||
Db    559  LMDTTWRDAHQSLLATRVRTHDLATIAPTTAHALAGRFALECWGGATFDVAMRFLHEDPW  618

Qy    596  DRLDELREAMPNVNIQMLLRGRNTVGTTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVS  655
                |||:||   | | |:|| | | |:   |  |||  |  || ::  ||||||||||||:|
Db    619  ERLRKLRSLVPNIPFQMLLRGANGVAYSSLPDNAIDHFVKQAKDNGVDIFRVFDALNDLE  678

Qy    656  QMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDM  715
                :  || :   ||||:| |  ||   :||:|  | |  | |||  :||   |:|| |||||
Db    679  QLKVGVDAVKKAG-GVVEATVCFSGDMLQPGKK-YNLDYYLEIIAEKIVQMGTHILGIKDM  736

Qy    716  AGLLRPAAVTKLVTALRREF-DLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSG  774
                ||  ||||||||   :  ::  ||:|||||  |  |:|  |||:|||||||  :  :||
Db    737  AGTMKPAAAKLLIGSLRAKYPDLPIHVHTHDSAGTAVASMTACALAGADVVDVAINSMSG  796

Qy    775  TTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLIYLPFESGTPGFTGRVVRHE  834
                || |||:|||    ||:|||  |||:|  | ||  |  || ||  || :| |||  |::|
Db    797  LTSQPSINALLASL-EGNIDTGINVEHVRELDAYWAEMRLLYSCFEADLKGPDPEVYQHE  855

Qy    835  IPGGQLSNLRAQATALGLLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAG  894
                ||||||:|||| ||  | : ||||  |   | |:   |  ||||:||||||| : :|||
Db    856  IPGGQLINLLFQAQQLGLGEQWAETKRAYREANYLLGDIVKVTPTSKVVGDLAQFMVSNK  915

Qy    895  VDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEE  954
                :     |||  |  |  :  |  :  |: |   ||   :|| | |   :     |  |
Db    916  LTSDDVRRLANSLDFPDSVMDFFEGLIGQPYGGFPEPFRSDVLRNK---RRKLTCRPGLE  972

Qy    955  QAHLDADDSKERRNSLNRL----------LFPKPTEEFLEHRRRFGNTSALDDREFFY  1002
                       :|   | |  :           : |  ||  | | | ||  ||: | |::
Db-   973  LEPFDLE--KIREDLQNRFGDVDECDVASYNMYPRVYEDFQKMRETYGDLSVLPTRSFLS  1030

Qy   1003  GLVEGRETLIRLPDVRTPLLVRLDAISEPDDK-GMRNVVANVNGQIRPMRVRDRSVESVT  1061
                 |:::|:||    :|    |   :| |  ::|    |::  :: || |  ||  | :| |
Db   1031  PLETDEEIEVVIEQGKT-LIIKLQAVGDLNKKTGEREVYFDLNGEMRKIRVADRSQKVET  1089

Qy   1062  ATAEKADSSNKGHVAAPFAGV-VTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKID  1120
                 |:|| |  |     |  ||   : |: | |  :   ||| |  |||:  || ::|| |
Db   1090  VTKSKADMHDPLHIGAPMAGVIVEVKVKHGSLIKKGQPVAVLSAMKMEMIISPSDGQVK  1149

Qy   1121  RVVVPAATKVEGGDLIVVV  1139
                |  :|     ||::| |:
Db   1150  EVFVSDGENVDSSDLLVLL  1168
```

Search completed: September 24, 2003, 16:00:32
Job time : 39.741 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      September 24, 2003, 15:44:03 ; Search time 18.3631 Seconds
                                 (without alignments)
                                 2919.469 Million cell updates/sec

Title:          US-09-974-973A-19
Perfect score:  5788
Sequence:       1 MSTHTSSTLPAFKKILVANR.........RVVVPAATKVEGDLIVVVS 1140

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters:          127863

Minimum DB seq length:  0
Maximum DB seq length:  2000000000

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :   SwissProt_41:*

   Pred. No. is the number of results predicted by chance to have a
   score greater than or equal to the score of the result being printed,
   and is derived by analysis of the total score distribution.

                              SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2541.5	43.9	1178	1	PYC_MOUSE	Q05920 mus musculu
2	2525.5	43.6	1178	1	PYC_RAT	P52873 rattus norv
3	2520.5	43.5	1178	1	PYC_HUMAN	P11498 homo sapien
4	2457	42.4	1178	1	PYC1_YEAST	P11154 saccharomyc
5	2436	42.1	1180	1	PYC2_YEAST	P32327 saccharomyc
6	2391	41.3	1189	1	PYC_PICPA	P78992 pichia past
7	1079	18.6	501	1	PYCA_METJA	Q58626 methanococc
8	1036.5	17.9	447	1	ACCC_ANASP	Q06862 anabaena sp
9	1019	17.6	506	1	PYCA_ARCFU	O30019 archaeoglob
10	997	17.2	491	1	PYCA_METTH	O27939 methanobact
11	946	16.3	448	1	ACCC_HAEIN	P43873 haemophilus
12	938	16.2	449	1	ACCC_ECOLI	P24182 escherichia
13	936	16.2	449	1	ACCC_ECO57	Q8X9b6 escherichia
14	924	16.0	449	1	ACCC_PSEAE	P37798 pseudomonas
15	911	15.7	725	1	MCCA_HUMAN	Q96rq3 homo sapien
16	909	15.7	717	1	MCCA_MOUSE	Q99mr8 mus musculu
17	883.5	15.3	654	1	BCCA_MYCTU	P46401 mycobacteri
18	879	15.2	448	1	ACCC_BACSU	P49787 bacillus su
19	875.5	15.1	598	1	BCCA_MYCLE	P46392 mycobacteri
20	871.5	15.1	567	1	PYCB_METJA	Q58628 methanococc
21	866	15.0	703	1	PCCA_HUMAN	P05165 homo sapien
22	859.5	14.8	731	1	PCCA_SOYBN	Q42777 glycine max
23	859	14.8	734	1	MCCA_ARATH	Q42523 arabidopsis
24	843.5	14.6	704	1	PCCA_RAT	P14882 rattus norv
25	835.5	14.4	1835	1	DUR1_YEAST	P32528 saccharomyc
26	825	14.3	568	1	PYCB_METTH	O27179 methanobact
27	804	13.9	590	1	DCOA_SALTY	Q03030 salmonella
28	799.5	13.8	595	1	DCOA_KLEPN	P13187 klebsiella
29	612.5	10.6	2345	1	COA1_RAT	P11497 rattus norv
30	609.5	10.5	2324	1	COA1_CHICK	P11029 gallus gall
31	609.5	10.5	2346	1	COA1_HUMAN	Q13085 homo sapien
32	608.5	10.5	2346	1	COA1_BOVIN	Q9ts3 bos taurus
33	601	10.4	2346	1	COA1_SHEEP	Q28559 ovis aries
34	561.5	9.7	2483	1	COA2_HUMAN	O00763 homo sapien
35	556.5	9.6	2273	1	HFA1_YEAST	P32874 saccharomyc
36	537.5	9.3	2233	1	COAC_YEAST	Q00955 saccharomyc
37	511	8.8	2280	1	COAC_SCHPO	P78820 schizosacch
38	267.5	4.6	1076	1	CARB_ARCFU	O28994 archaeoglob
39	264	4.6	1060	1	CARB_METTH	O27077 methanobact
40	255	4.4	1072	1	CARB_ECO57	Q8xa38 escherichia
41	255	4.4	1072	1	CARB_ECOLI	P00968 escherichia
42	254.5	4.4	1077	1	CARB_VIBVU	Q8dem2 vibrio vuln
43	253.5	4.4	1099	1	CARB_THEMA	Q9wz27 thermotoga
44	251.5	4.3	1074	1	CARB_SALTY	P14846 salmonella
45	251.5	4.3	1077	1	CARB_VIBPA	Q87sf3 vibrio para

```
                              ALIGNMENTS

RESULT 1
PYC_MOUSE
ID   PYC_MOUSE      STANDARD;      PRT;   1178 AA.
AC   Q05920;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic
DE   carboxylase) (PCB).
GN   PC OR PCX.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Adipocyte;
RX   MEDLINE=93189578; PubMed=8446588;
RA   Zhang J., Xia W.L., Brew K., Ahmad F.;
RT   "Adipose pyruvate carboxylase: amino acid sequence and domain
RT   structure deduced from cDNA sequencing.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:1766-1770(1993).
CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
CC       CARBOXYL GROUP TO PYRUVATE IN THE SECOND. CATALYZES IN A TISSUE
CC       SPECIFIC MANNER, THE INITIAL REACTIONS OF GLUCOSE (LIVER, KIDNEY)
CC       AND LIPID (ADIPOSE TISSUE, LIVER, BRAIN) SYNTHESIS FROM PYRUVATE.
CC   -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
CC       oxaloacetate.
CC   -!- COFACTOR: BIOTIN AND MANGANESE.
CC   -!- PATHWAY: GLUCONEOGENESIS AND LIPOGENESIS.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC   -!- TISSUE SPECIFICITY: LIVER, KIDNEY, ADIPOSE TISSUE, LIVER
CC       AND BRAIN.
CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC       AND CARBAMYL PHOSPHATE SYNTHETASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L09192; AAA39737.1; -.
DR   PIR; A47255; A47255.
DR   HSSP; P24182; 1BNC.
DR   SWISS-2DPAGE; Q05920; MOUSE.
DR   MGD; MGI:97520; Pcx.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
```

```
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_Carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
KW   Ligase; Multifunctional enzyme; Biotin; Manganese; Gluconeogenesis;
KW   ATP-binding; Mitochondrion; Lipid synthesis; Transit peptide.
FT   TRANSIT       1     20       MITOCHONDRION (POTENTIAL).
FT   CHAIN        21   1178       PYRUVATE CARBOXYLASE.
FT   DOMAIN       21    549       BIOTIN CARBOXLASE (BY SIMILARITY).
FT   DOMAIN      550   1000       CARBOXYLTRANSFERASE (BY SIMILARITY).
FT   DOMAIN     1096   1178       BIOTIN CARBOXYL CARRIER PROTEIN
FT                                (BY SIMILARITY).
FT   NP_BIND     198    203       ATP (BY SIMILARITY).
FT   ACT_SITE    328    328       BY SIMILARITY.
FT   BINDING    1144   1144       BIOTIN (BY SIMILARITY).
SQ   SEQUENCE   1178 AA;  129684 MW;  14CEAUF9DA8B8127 CRC64;

Query Match             43.9%;  Score 2541.5;  DB 1;  Length 1178;
Best Local Similarity   47.0%;  Pred. No. 6.7e-129;
Matches  537;  Conservative 187;  Mismatches 401;  Indels  17;  Gaps  11;

Qy      13 KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL   72
           ::::|||||||||| || |  |:|    |||::|| : : :|::|::|  | ::|::|||
Db      38 KKVMVANRGEIAIRVFRACTELGIRTVAVYSEQDTGQMHRQKADEAYLIGRGLAPVQAYL   97

Qy      73 DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT  132
           ::: ||| ::  :|  |:: ||:||:|  |  |  ||:|:|||||:|||::::|:||
Db      98 HIPDIIKVAKENGVDAVHPGYGFLSERADFAQACQDAGVRFIGPSPEVVRKMGDKVEARA  157

Qy     133 AAKKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKLA  191
            :||:| |::  || || | :   : || ||:|:||||:|:|||||||||||||:|||:|
Db     158 IAIAAGVPVVPGTDSPISSLHEAHEFSNTFGFPIIFKAAYGGGGRGMRVVHSYEELEENY  217

Qy     192 TEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVEI  251
           || | | |||||||||||| :  ||| |::|||||:||:||:|||||:|:|||||||||
Db     218 TRAYSEALAAFGNGALFVEKFIEKPRHIEVQILGDQYGNILHLYERDCSIQRRHQKVVEI  277

Qy     252 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE  311
           ||| |||| || ::: || ||| | :| :|||||||:|| |:|||||:|||:|||:|||:
Db     278 APATHLDPQLRSRLTSDSVKLAKQVGYENAGTVEFLVDKHGKHYFIEVNSRLQVEHTVTE  337

Qy     312 EVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAY  371
           |:||:|||:||| ||||:|  | :|  |:|   |  |:||:|::||| :|||||:||:||
Db     338 EITDVDLVHAQIHVSEGRSLPDLGLRQENIRINGCAIQCRVTTEDPARSFQPDTGRIEVF  397

Qy     372 RSPGGAGVRLDGAAQL-GGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT  430
           ||  |:|:|| ::|:|  |   | ||| :|| :  ::: |||:||||||||||||||||:
Db     398 RSGEGMGIRLDNASAFQGAVISPHYDSLLVKVIAHGKDHPTAATKMSRALAEFRVRGVKT  457

Qy     431 NIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVNKPHGVRP  490
           ||||:|| | | | :|::|||| ||||||:   | |||:| :|:|||||:|||||||:||
Db     458 NIPFLQNVLNNQQFLAGTVDTQFIDENPELFQLRPAQNRAQKLLHYLGHVMVNGPTTPIP  517

Qy     491 KDVA-APID-KLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVIDTTFRDAHQSL  548
           :::   ||:  ||:|||  ||| : :|:|||| ||:| ::||||| |:||||||||||||
Db     518 VNVSPSPVDPAVPVVPIGPPPAGFRDILLREGPEGFARAVRNHQGLLLMDTTFRDAHQSL  577

Qy     549 LATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNV  608
           ||||||:|  | |  |  |:  | :: ||:|||:||||||:|||||||||::||  :|::
Db     578 LATRVRTHDLKKIAPYVAHNFNKLFSMENWGGATFDVAMRFLYECPWRRLQELRELIPNI  637

Qy     609 NIQMLLRGRNTVGTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETN  668
           ::|:||||:|  | : |:||| |||  : : |||  :||| ||  : :| ||:|  :  |
Db     638 PFQMLLRGANAVGYTNYPDNVVKFCEVAKENGMDVFRVFDSLNYLPNMLLGMEAAGSAG  697
```

```
Qy     669 TAVAEVAMAYSGDLSDPNEKLYTLDYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLV  728
               | :|:::||: ||: |:  ||||||  :|||:||:||:|||||||||| :::||||
Db     698 -GVEEAISYTGDVADPSRTKYSLEYYMGLAEELVRAGTHILCIKDMAGLLKPAACTMLV  756

Qy     729 TALRREF-DLPVHVHTHDTAGGQLAIYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAA  787
              |||:|  ||: |:|||||| |  :  : :||||| ||  |    :: |: |: |: |
Db     757 SSLRDRFPDLPLHIHTHDTSGAGVAAMLACAQAGADVVDVSMSGMTSQPSMGALVAC  816

Qy     788 FAHTRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGR--VIRHEIPGGQLSNLRA  845
            |||:|||| ||: | ||| |  |:|||| | ||:||||::    || |||| :||:|:
Db     817 TKGTPLDTEVPLERVFDYSEYWEGARGLYAAFDCTATMKSGNSDVYENEIPGGQYINLHF  876

Qy     846 QATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAADPQ  905
           ||: |||:|:||:::|  :::: ||| |::|||||||:||| ::|:: | |  | : |
Db     877 QAHSMGLGSKPKEVKKAYVEANQMLGDLIKVTPSSKIVGDLAQFMVQNGLSRAEAEAQAE  936

Qy     906 KYDIPDSVIAFLRGELGNPPGGWPEPLRTRAL------EGRSEGKAPLTEVPEEEQAHLD  959
            |: | ::    ::  | :|||:  |::::::       |  : || |      |: : |
Db     937 ELSFPRSVVEFLQGYIGIPHGGFPEPFRSKVLKDLPRIEGRPGASLPPLNKLELEKDLID  996

Qy     960 A-DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVR 1018
             |::::|| :|:||::  | ::::| ::|:|:|:  : ::    |:|:|: :|:|: :|:
Db     997 RHGEEVTPEDVLSAAMYPDVFAQFKDFTATFGPLDSLNTRLFLQGPKIAEEFEVELERGK 1056

Qy    1019 TPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAP 1078
            | | ||  |:||||:| ::   :|   :||:|::|| : |::|   ||||:|:||| ||
Db    1057 T-LHIKALAVSDLNRAGQRQVFFELNGQLRSILVKDTQAMKEMHFHPKALKDVKGQIGAP 1115

Qy    1079 FAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIV 1137
           |||  |||  ||| ||| | : |::|||||||:| :|  ||  | ||::|||:| | |::
Db    1116 MPGKVIDIKVAAGDKVAKGQPLCVLSAMKMETVVTSPMEGTIRKVHVTKDMTLEGDDLIL 1175

Qy    1138 VV 1139
              ::
Db    1176 EI 1177
```

```
RESULT 2
PYC_RAT
ID   PYC_RAT      STANDARD;      PRT;   1178 AA.
AC   P52873; Q64555;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic
DE   carboxylase) (PCB).
GN   PC.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=96096548; PubMed=8522203;
RA   Lehn D.A., Moran S.M., Macdonald M.J.;
RT   "The sequence of the rat pyruvate carboxylase-encoding cDNA.";
RL   Gene 165:331-332(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Liver;
RX   MEDLINE=96257760; PubMed=8687410;
RA   Jitrapakdee S., Booker G.W., Cassady A.I., Wallace J.C.;
RT   "Cloning, sequencing and expression of rat liver pyruvate
RT   carboxylase.";
RL   Biochem. J. 316:631-637(1996).
CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
CC       CARBOXYL GROUP TO PYRUVATE IN THE SECOND. CATALYZES THE INITIAL
CC       SPECIFIC MANNER, THE INITIAL REACTIONS OF GLUCOSE (LIVER, KIDNEY)
CC       AND LIPID (ADIPOSE TISSUE, LIVER, BRAIN) SYNTHESIS FROM PYRUVATE.
CC   -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
```

```
CC        oxaloacetate.
CC   -!- COFACTOR: BIOTIN AND MANGANESE (BY SIMILARITY).
CC   -!- PATHWAY: GLUCONEOGENESIS AND LIPOGENESIS.
CC   -!- SUBUNIT: Homotetramer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC        AND CARBAMYL PHOSPHATE SYNTHETASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U32314; AAA96256.1; -.
DR   EMBL; U36585; AAC52668.1; -.
DR   PIR; S68252; JC4391.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
KW   Ligase; Multifunctional enzyme; Biotin; Manganese; Gluconeogenesis;
KW   ATP-binding; Mitochondrion; Lipid synthesis; Transit peptide.
FT   TRANSIT       1     20       MITOCHONDRION (POTENTIAL).
FT   CHAIN        21   1178       PYRUVATE CARBOXYLASE.
FT   DOMAIN       21    549       BIOTIN CARBOXYLASE (BY SIMILARITY).
FT   DOMAIN      550   1000       CARBOXYLTRANSFERASE (BY SIMILARITY).
FT   DOMAIN     1096   1178       BIOTIN CARBOXYL CARRIER PROTEIN
FT                                (BY SIMILARITY).
FT   NP_BIND     198    203       ATP (BY SIMILARITY).
FT   ACT_SITE    328    328       BY SIMILARITY.
FT   BINDING    1144   1144       BIOTIN (BY SIMILARITY).
FT   CONFLICT    222    222       P -> S (IN REF. 2).
FT   CONFLICT    222    222       D -> I (IN REF. 2).
FT   CONFLICT    866    866       G -> R (IN REF. 2).
SQ   SEQUENCE   1178 AA;  129689 MW;  8E5FA19BC132A8DD CRC64;
```

```
Query Match             43.6%;  Score 2525.5;  DB 1;  Length 1178;
Best Local Similarity   46.4%;  Pred. No. 4.8e-128;
Matches  541; Conservative 185; Mismatches 410; Indels 31; Gaps 13;

Qy         1 MSTHTSSTLPA----------FKKILVANRGEIAVRAFRAALETGAATVAIIYPREDRG 48
                                    :||||||||||:| || ||||:||| :|:|
Db        14 LGVRRSSTAPVASPNVRRLEYKPIKKWMVANRGEIAIRVFRACTELGIRTVAVYSEQDTG 73

Qy        49 SFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAE 108
                ::||::|| ||||:||  | |||::||:|| ||:||||:|::|||||||  :|| : |
Db        74 QMHRQKADEAYLIGRGLAPVQAYIHIPDIIKVAKENGVDAVHPGYGFLSERADFAQACQD 133

Qy       109 NGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIFV 167
                |:||||:|:||:|||| ||:|| :|||||||: :|| :||  ::|| :|   |:| :
Db       134 AGVRFIGPSPEVVRKMGDKVEARAIAAGVPVVPGINSPINSLHEAHEFSNTYGFPIIF 193

Qy       168 KAVAGGGGRGMRFVASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDH 227
                || |||||||| |:|||::|| | ||: ||:|||:|||||: ||||:||||||||||
Db       194 KAAYGGGGRGMRVVHSYEELEENYTRAYPEALAAFGNGALFVEKFIEKPRHIEVQILGDQ 253

Qy       228 TGEVVHLYERDCSLQRRHQKVVEIAPAQHLDEPELRDRICADAVKFCRSIGYQGAGTVEFL 287
```

```
Db       254 YGNILHLYERDCSIQRRHQKVVEIAPATHLDPQLRSRLTSDSVKLAKQVGYENAGTVEFL 313
                 ||:|||||||||||||||||||||||||| ::||:|:||:||:|| :|:|:||||||||
Qy       288 VDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKEIGLTQDKIKTHGAA 347
                 |:| ||:||||||||:|||::||||:|| ::: : |: |:|| ::| ::|| :||:|:|
Db       314 VDKHGKHYFIEVNSRLQVEHTVTEEITDVDLVHAQIHVSEGRSLPDIGLRQENIRINGCA 373
Qy       348 LQCRITTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQL-GGEITAHFDSMLVKMTCRG 406
                 :||::||||:||||||||||:||||||||||:|||:::|  ||||:||||:|:|:|| |
Db       374 IQCRVTTEDPARSFQPDTGRIEVFRSGEGMGIRLDNASAFQGAVISPHYDSLLVKVIAHG 433
Qy       407 SDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKKRIATGFIADRPHLLQAPPA 466
                 |||||| | :|||:|| ||| || || | |: ||:|: || :| ||| || |:| :|:
Db       434 KDHPTAATKMSRALAEFRVRGVKTNIPFLQNVLNNQQFLAGIVDTQFIDENPELFQLRPA 493
Qy       467 DDEQGRILDYLADVTVNKPHGVRPKDVA-APIDKL-PNIKDLPLPRGSRDRLKQLGPAAF 524
                 ::: | || ||||:|| ||||||||||   || :  |: :|:|||::|||::|:||:||
Db       494 QNRAQKLLHYLGHVMVNGPTTPIPVKVSPSPVDPIVPVVPIGPPAGFRDILLREGPEGF 553
Qy       525 ARDLREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYD 584
                 || :|:|| || |::|:||||| |||||||| : |||: :||::| ||:|| |||||:|
Db       554 ARAVRNHQGLLLMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFNNLFSIENWGGATFD 613
Qy       585 VAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDI 644
                 ||||| || || ||:||:|| :||:: ||||| || |||||:|||||:::|::::||:
Db       614 VAMRFLYECPWRRLQELRELIPNIFQMLLRGANVGYTNYPDNVVFKFCEVAKENGMDV 673
Qy       645 FRIPDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYIKMAAEIVK 704
                 |||| ||:|:|||: ::|| |: :|| || || :||::| : || |:| ::|:|||:||
Db       674 FRIFDSLNYLPNMLLGMEAARGSAG-GVVEAAISYTGDVADPSRTKYXSLEYYMGLAEELVR 732
Qy       705 SGAHILAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVHTHDTAGGQLATYFAAAQAGAD 763
                 || ||:||||| |:|||| |||:|:|||::  :||||:||||:|||| |||||||||||
Db       733 AGTHILCIKDMAGLLKPAACTMLVSSLRDRFPDLPLHIHTHDTSGSGVAAMLACAQAGAD 792
Qy       764 AVDGASAPLSGTTSQPSLSALVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGT 823
                 | |||| || ||:||||||:||:||||||||||:|||: |||||:|| ||||:||||||
Db       793 VVDVAVDSMSGMTSQPSMGALVACTKGTPLDTEVPLERVFDYSEYWEGARGLYAAFDCTA 852
Qy       824 PGPTGR--VYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSK 881
                 ||| |   |:| |||| |:|:||:| ||:||||| :| :||| :::::|| |||||||||
Db       853 TMKSGNSDVYENEDPGGQYINLHFQAHSMGLGSKFEVKKAYVEANQMLGDLIKYTPSSK 912
Qy       882 VVGDIALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEG-- 939
                 :::|: :|| || || ||: |||:|||:|:|::||||||:|||||||||||:|||:||
Db       913 IVGDLAQFMVQNGLSRAEAEAQAEELSFPRSVVEFLQGYIGIPHGGFPEPFRSKVLKDLP 972
Qy       940 RSEGKRAPLTEVPEEEQAHLDAD------DSKERRNLRLLFKPTEEFLEHRRRFGNTS 993
                 | ::  |||  || ::|| || |      : |||:|||||||:|||||:||||| |:|
Db       973 RIEG-GPGASLLPPLNLKELEKDLIDRHGEEVTPEDVLSAAMYPDVFAQFKDFTATFGPLD 1031
Qy       994 ALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVR 1053
                 :|::|::||:||:||: ||||| ||:| :||||||:: : :| |:||::|:|: ||:||
Db      1032 SLNTRLFLQGFKIAEEFEVELERGKT-LHIKALAVSDLNRAGQRQYFFELNGQLRSIIVK 1090
Qy      1054 DRSVESVTATAEKADSSNKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATIT 1112
                 :|: | | |:: :::  |||:||||:|| |:::|:||||||||| || |: |:|| ||:|
Db      1091 DTQAMKEMHFHPKALKDVKGQIGAPMPGKVIDVKVAAGAKVVKGQPLCVLSAMKMETVVT 1150
Qy      1113 ASVDGKIDRVVPAATKVEGDLIVVV 1139
                 :| |:|| |||::||:| || |:
Db      1151 SPMEGTIRKVHVTKDMTLEGDDLILEI 1177
```

```
RESULT 3
PYC_HUMAN
ID   PYC_HUMAN      STANDARD;      PRT;   1178 AA.
AC   P11498; Q16705;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic
DE   carboxylase) (PCB).
```

```
GN   PC.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney, and Liver;
RX   MEDLINE=95002202; PubMed=7918683;
RA   Wexler I.D., Du Y., Lisgaris M.V., Mandal S.K., Freytag S.O.,
RA   Yang B.-S., Liu T.-C., Kwon M., Patel M.S., Kerr D.S.;
RT   "Primary amino acid sequence and structure of human pyruvate
RT   carboxylase.";
RL   Biochim. Biophys. Acta 1227:46-52(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=94324922; PubMed=8048912;
RA   Mackay N., Rigat B., Douglas C., Chen H.S., Robinson B.H.;
RT   "cDNA cloning of human kidney pyruvate carboxylase.";
RL   Biochem. Biophys. Res. Commun. 202:1009-1014(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney, and Liver;
RA   Walker M.E., Jitrapakdee S., Val D.L., Wallace J.C.;
RL   Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 1083-1178 FROM N.A.
RX   MEDLINE=87212051; PubMed=355348;
RA   Lamhonwah A.-M., Quan F., Gravel R.A.;
RT   "Sequence homology around the biotin-binding site of human propionyl-
RT   CoA carboxylase and pyruvate carboxylase.";
RL   Arch. Biochem. Biophys. 254:631-636(1987).
RN   [6]
RP   SEQUENCE OF 1135-1178 FROM N.A.
RX   MEDLINE=85030380; PubMed=6548474;
RA   Freytag S.O., Collier K.J.;
RT   "Molecular cloning of a cDNA for human pyruvate carboxylase.
RT   structural relationship to other biotin-containing carboxylases and
RT   regulation of mRNA content in differentiating preadipocytes.";
RL   J. Biol. Chem. 259:12831-12837(1984).
RN   [7]
RP   VARIANTS PC DEFICIENCY THR-610 AND ILE-743.
RX   MEDLINE=98254451; PubMed=9585512;
RA   Carbone M.A., MacKay N., Ling M., Cole D.E.C., Douglas C., Rigat B.,
RA   Feigenbaum A., Clarke J.T.R., Haworth J.C., Greenberg C.R.,
RA   Seargeant L., Robinson B.H.;
RT   "Amerindian pyruvate carboxylase deficiency is associated with two
RT   distinct missense mutations.";
RL   Am. J. Hum. Genet. 62:1312-1319(1998).
RN   [8]
RP   VARIANTS PC DEFICIENCY ALA-145 AND CYS-451.
RX   MEDLINE=9824401; PubMed=9585002;
RA   Wexler I.D., Kerr D.S., Du Y., Kaung M.M., Stephenson W., Lusk M.M.,
RA   Wappner R.S., Higgins J.J.;
RT   "Molecular characterization of pyruvate carboxylase deficiency in two
RT   consanguineous families.";
RL   Pediatr. Res. 43:579-584(1998).
CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
CC       CARBOXYL GROUP TO PYRUVATE IN THE SECOND. CATALYZES IN A TISSUE
CC       SPECIFIC MANNER, THE INITIAL REACTIONS OF GLUCOSE (LIVER, KIDNEY)
CC       AND LIPID (ADIPOSE TISSUE, LIVER, BRAIN) SYNTHESIS FROM PYRUVATE.
CC   -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
CC       oxaloacetate.
CC   -!- COFACTOR: BIOTIN AND MANGANESE.
CC   -!- PATHWAY: GLUCONEOGENESIS AND LIPOGENESIS.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC   -!- DISEASE: Defects in PC are the cause of pyruvate carboxylase
CC       deficiency (PC deficiency) [MIM:266150] leading to lactic
CC       acidosis, mental retardation and death. It occurs in three forms:
CC       mild or type A, severe neonatal or type B, and a very mild
CC       lacticacidemia.
CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC       AND CARBAMYL PHOSPHATE SYNTHETASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U04641; AAA99537.1; -.
DR   EMBL; S72370; AAB31500.1; -.
DR   EMBL; U30891; AAA82937.1; -.
DR   EMBL; BC011617; AAH11617.1; -.
DR   EMBL; M26122; AAA36423.1; -.
DR   EMBL; K02282; AAA60033.1; -.
DR   PIR; G01933; JC2460.
DR   HSSP; P24182; 1BNC.
DR   Genew; HGNC:8636; PC.
DR   GK; P11498; -.
DR   MIM; 266150; -.
DR   GO; GO:0005524; F:ATP binding activity; TAS.
DR   GO; GO:0009374; F:biotin binding activity; TAS.
DR   GO; GO:0004736; F:pyruvate carboxylase activity; TAS.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; Biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
KW   Ligase; Multifunctional enzyme; Biotin; Manganese; Gluconeogenesis;
KW   ATP-binding; Mitochondrion; Lipid synthesis; Transit peptide;
KW   Disease mutation.
FT   TRANSIT       1     20       MITOCHONDRION (POTENTIAL).
FT   CHAIN        21   1178       PYRUVATE CARBOXYLASE.
FT   DOMAIN       21    549       BIOTIN CARBOXYLASE (BY SIMILARITY).
```

```
FT   DOMAIN        550   1000       CARBOXYLTRANSFERASE (BY SIMILARITY).
FT   DOMAIN       1096   1178       BIOTIN CARBOXYL CARRIER PROTEIN
FT                                  (BY SIMILARITY).
FT   NP_BIND       198    203       ATP (BY SIMILARITY).
FT   ACT_SITE      328    328       BY SIMILARITY.
FT   BINDING      1144   1144       BIOTIN (BY SIMILARITY).
FT   VARIANT       145    145       V -> A (in PC deficiency; mild).
FT                                  /FTId=VAR_015199.
FT   VARIANT       451    451       R -> C (in PC deficiency; mild).
FT                                  /FTId=VAR_015200.
FT   VARIANT       610    610       A -> T (in PC deficiency; mild).
FT                                  /FTId=VAR_008095.
FT   VARIANT       743    743       M -> I (in PC deficiency; mild).
FT                                  /FTId=VAR_008096.
FT   CONFLICT      225    226       LA -> WP (IN REF. 2).
FT   CONFLICT      352    352       A -> S (IN REF. 3).
FT   CONFLICT      385    386       RS -> PT (IN REF. 2).
FT   CONFLICT      486    487       EL -> DV (IN REF. 2).
FT   CONFLICT      638    638       P -> R (IN REF. 2).
FT   CONFLICT      729    729       E -> A (IN REF. 2).
FT   CONFLICT      774    775       DT -> AP (IN REF. 2).
SQ   SEQUENCE     1178 AA;   129633 MW;  381F527553A20095 CRC64;
```

```
Query Match             43.5%;
Best Local Similarity   46.3%;   Score 2520.5;  DB 1;  Length 1178;
Matches  537;  Conservative 183;  Mismatches 412;  Indels  29;  Gaps  12;

Qy      6  SSTLPA----------FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRS  53
              ::|                ||:|||||||:| ||| |||||| |: ||     ||
Db     19  TSTPAASPNVRRLEYKPIKKVMVANRGEIAIRVFRACTELGIRTVAIYSEQDTGQMHRQ  78

Qy     54  FASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAITPGYGFLSENAQLARECAENGITF  113
           |:|| ||| |:  ||:  ||||:||:|:||:| ::|||||||| ||||   | |:|||||
Db     79  KADEAYLIGRGLAPVQAYLHIPDIIKVAKENNVDAVHPGYGFLSERADFAQACQDAGVRF  138

Qy    114  IGPTPEVLDLTGDKSRAVTAAKKAGLPVL-AESTPSKNIDEIVKSAEGQTYIFVKAVAG  172
           ||:||::|:   ||| |:|   ||||| :   :::|   :|:||||| | ||||||||:
Db    139  IGPSPEVVRKMGDKVEARAIAIIAAGVPVPGTDAPITSLHEAHEFSNTYGFPIIFKAAYG  198

Qy    173  GGGRGMRFVASPDELRKLATEASREARAAFGDGAVYVERAVINPQHIEVQILGDHTGEVV  232
           |||||||| ::||:|||| |  ||| ||| ||| |:|:| ||| |||:|||||| ||:
Db    199  GGGRGMRVHSYELEENYTRAYSEALAAFGNGALFVEKFIEKPRHIEVQILGDQYGNIL  258

Qy    233  HLYERDCSLQRRHQKVVELAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKG  292
           |||||| | |||||||||||||||| |||| || ||| |||:||||||  ||||||||:|
Db    259  HLYERDCSIQRRHQKVVELAPAAHLDPQLRTRLTSDSVKLAKQVGYENAGTVEFLVDRHG  318

Qy    293  NHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRI  352
            ||:||:||||:|| | |:|  ||:||||| |||: ||| |: | ::||| ||||| ||
Db    319  KHYFIEVNSRLQVEHTVTEEITDVDLVHAQIHVAEGRSLPDLGLRQENIRINGCAIQCRV  378

Qy    353  TTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQL-GEEITAHFDSMLVKMTCRGSDFET  411
           |||| ||||| ||| :| :|:| || ||||  :|   ||: ||:|: ||: |||: ||:
Db    379  TTEDPARSFQPDTGRIEVFRSGEGMGIRLDNASAFQGAVISPHYDSLLVKVIAHGKDHPT  438

Qy    412  AVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGTPIADHPHLLQAPPADDEQG  471
           | ||||  | || |||| ||| || || ||||  ||  :| |:||  ::|||  |  ::|
Db    439  AATKMSRALAEFRVRGVKTNIAFLQNVLNNQQFLAGTVDTQFIDENPELFQLRPAQNRAQ  498

Qy    472  RILDYLADVTVNKPHGVRP-KDVAAPIDK-LPNIKDLPLPRGSRDRLKQLGPAAFARDLR  529
           ||||| ||||:||||| |  ||:||||:|  |||| ||| || || || ||||| |:|::
Db    499  KLLHYLGHVMVNGPTTPIPVKASPSPTDPVVPAVPIGPPAGFRDILLREGPEGFARAVR  558

Qy    530  EQDALAVDTDTTFRDAHQSLLATPVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRF  589
           | | ||| |  ||||||||| |  :|:  |  | |::  || |:|:|| |||| ||| ||
Db    559  NHPGLLLMDTTFRDAHQSLLATPVRTHDLKKIAPYVAHNFSKLFSMENWGGATFDVAMRF  618

Qy    590  LFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFD  649
           ||   |||::|| : |||||:|||||| || ||| ||| ||  ||||| |||||| |||
Db    619  LYECPWPRRLQELRELIPNIPFQMLLRGANAVGYTNYPDNVVFKFCEVAKENGMDVFRVFD  678

Qy    650  ALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYILKMAEELVKSGAHI  709
           :||  ::    ||   |:   :   :|: |||  ||  ||| |||| | ::|: |:  ||
Db    679  SLNYLPNMLLGMEAAGSAG-GVEEAAISYTGDVADPSRTKYSLQYYMGLAEELVRAGTHI  737

Qy    710  LAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVHTHDTAGGQLATYFAAAQAGADAVDGA  768
           |||||||||| | | |:|:|||:|:|  ||||:||||||| ||| |:|:||||||| ||
Db    738  LCIKDMAGLLKPTACTMLVSSLRDRFPDLPLHIHTHDTSGAGVAAMLACAQAGADVVDVA  797

Qy    769  SAPLSGTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFFESGTGPTG  828
           ::   |||  |  :::   | | | ||  :  :  |  :::  |||  ||:|| ||:|
Db    798  ADSMSGMTSQPSMGALVACTRGTPLDTEVPMERVFDYSEYWEGARGLYAAFDCTATMKSG  857

Qy    829  R--VYRHEIPGGQLSNLRAQATALGLADREFLIEDNYAAVNEMLGRPTKVTPSSKVVGDL  886
           |   | | ||||| :    | | | | :  |: ::   :  ::||    |||:|| ||||
Db    858  NSDVYENEIPGGQYTNHFQAHSMGLGSKFKEVKKAYVEANQMLGDLIKVTPSSKIVGDL  917

Qy    887  ALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRAL-----EGR  940
           |  |||:| ||   | ||  |::|:|::|: |:||  | || | |: |::|      |
Db    918  AQFMVQNGLSRAEAPAQAEELSFPRSVVEFLQGYIGVPHGGFPEPFRSKVLKDLPRVEGR  977

Qy    941  SEGKAPLTEVPEEQAHLDA-DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDRE  999
            ::   :  :| : : |:   : :: ::  |   |  | ::  |: |:| |:| |:|
Db    978  PGASLPPLDLQALEKELVDRHGEEVTPEDVLSAAMYPDVPAHFKDFTATFGPLDSLNTRL  1037

Qy   1000  FFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNQGIRPMRVRDRSVES  1059
                 :|  | | | |  |     |: |   :  | | :: | :| |:|  | : :|::
Db   1038  FLQGPKIAEEFEVELERGKT-LHIKALAVSDLNRAGQRQVFFELNGQLRSIIVKDTQAMK  1096

Qy   1060  VTATAEKADSSNKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGK  1118
           ||                     |:|:|: ||| |||:||| :||||| : |||::|
Db   1097  EMHFHPKALKDVKGQIGAPMGPKVIDIKVVAGAKVAKGQPICVLSAMKMETVVISPMEGT  1156

Qy   1119  IDRVVVPAATKVEGGDLIVVV  1139
           : ::|   |||::||| |:::
Db   1157  VRKVHVTKDMTLEGDDLILEI  1177
```

```
RESULT 4
PYC1_YEAST
ID   PYC1_YEAST    STANDARD;      PRT;   1178 AA.
AC   P11154;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyruvate carboxylase 1 (EC 6.4.1.1) (Pyruvic carboxylase 1) (PCB 1).
GN   PYC1 OR PYV OR YGL062W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=88298805; PubMed=3042770;
RA   Lim F., Morris C.P., Occhiodoro F., Wallace J.C.;
RT   "Sequence and domain structure of yeast pyruvate carboxylase.";
RL   J. Biol. Chem. 263:11493-11497(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=97377993; PubMed=9234674;
RA   Feuermann M., de Montigny J., Potier S., Souciet J.-L.;
RT   "The characterization of two new clusters of duplicated genes
RT   suggests a 'Lego' organization of the yeast Saccharomyces cerevisiae
RT   chromosomes.";
RL   Yeast 13:861-869(1997).
RN   [3]
RP   SEQUENCE OF 1003-1178 FROM N.A.
RX   MEDLINE=87241529; PubMed=3036126;
RA   Morris C.P., Lim F., Wallace J.C.;
RT   "Yeast pyruvate carboxylase: gene isolation.";
RL   Biochem. Biophys. Res. Commun. 145:390-396(1987).
CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
CC       CARBOXYL GROUP TO PYRUVATE IN THE SECOND.
```

```
CC   -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
CC       oxaloacetate.
CC   -!- COFACTOR: BIOTIN AND ZINC.
CC   -!- PATHWAY: Gluconeogenesis.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC       AND CARBAMYL PHOSPHATE SYNTHETASES.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; J03889; AAA34843.1; -.
DR   EMBL; Z72584; CAA96765.1; -.
DR   PIR; S64066; QYBYP.
DR   HSSP; P24182; 1BNC.
DR   SGD; S0003030; PYC1.
DR   GO; GO:0005829; C:cytosol; IDA.
DR   GO; GO:0004736; F:pyruvate carboxylase activity; IDA.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Ligase; Multifunctional enzyme; Biotin; Gluconeogenesis; ATP-binding;
KW   Zinc; Multigene family.
FT   NP_BIND      182    187       ATP (POTENTIAL).
FT   ACT_SITE     312    312       BY SIMILARITY.
FT   BINDING     1135   1135       BIOTIN (BY SIMILARITY).
FT   CONFLICT     462    462       T -> G (IN REF. 1).
FT   CONFLICT     493    493       V -> D (IN REF. 1).
FT   CONFLICT     595    595       R -> A (IN REF. 1).
FT   CONFLICT     619    619       E -> Q (IN REF. 1).
FT   CONFLICT     664    664       G -> S (IN REF. 1).
FT   CONFLICT     772    772       A -> R (IN REF. 1).
FT   CONFLICT     879    879       E -> Q (IN REF. 1).
FT   CONFLICT     909    909       Q -> K (IN REF. 1).
SQ   SEQUENCE   1178 AA;  130099 MW;  BC7110ABAFB23E04 CRC64;

Query Match             42.4%;  Score 2457;  DB 1;  Length 1178;
Best Local Similarity   46.7%;  Pred. No. 2.3e-124;
Matches  541;  Conservative  173;  Mismatches  401;  Indels  44;  Gaps  16;

Qy    14 KILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSPASEAVRIGTEG--SPVKAY 71
          ||||||||||  |  ||:|  ::  ||||| ||  |:  | |:||    |: : | :||||
Db    21 KILVANRGEIPIRIFRTAHELSMQTVAIYSHEDRLSTHKQKADEAYVIGEVGQYQYITPVGAY 80

Qy    72 LDIDEIIGAAKKVKADAIYPSYGFLSENAQLARECAENGITFIGPTPEVLLTGDKSRAV 131
          | |||| :||:|:|.|.|:|.||||||||:: |: ||| ||:|| .||:.:.|||.||
Db    81 LAIDEIISIAQKHQVDFIHPGYGFLSENSEFADKVVKAGITWIGPPAEVIDSVGDKVSAR 140

Qy   132 TAAKKAGLPVLAEST-PSKNIDEIVKSAEQGTYPIFVKAVAGGGGRGMRFVASPDELRKL 190
          .||||| |.|:||:. ||:||:|:|::||:|:|| |:|||.|||||||||::|  |:.|
Db   141 NLAAKANVFVPGPGPIETVEEAALDFVNEYGYPVIIKAAFGGGGRGMRVVREGDDVADA 200
```

```
Qy   191 ATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLLERDCSLQRRHQKVVE 250
          .|:|:||| :|||:|: :||| :|:| :||||:| |:.|.|||||||||:|||||||||||
Db   201 FQRATSEARTAFGNGTCFVERFLDKPKHIEVQLLADNHGNVVHLFERDCSVQRRHQKVVE 260

Qy   251 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 310
          :||| .|| |:||.|:.|||| |.::||:.||||||:||::.:|:||||:|||||||||:|
Db   261 VAPAKTLPREVRDALITDAVKLAKECGYRNAGTAEFLVDNQNRHYFIEINPRIQVEHTIT 320

Qy   311 EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 370
          ||:| :|:|.||::|||||:| :|||:|||||| | | |||||||||:.:|:|||| |
Db   321 EEITGIDIVAAQIOIAAGASLPQLGLFQDKITTRGFAIQCRITTEDPAKNFQPDTGRIEV 380

Qy   371 YRSPGGAGVRLDGA-AQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVA 429
          |||.||.||||||. | .|  ||:|:|||||| |  .|:: |:|.|::|.|:||::.||
Db   381 YRSAGGNGVRLDGGNAYAGTIISPHYDSMLVKCSCSGSTYEIVRKMIRALIEFRIRGVK 440

Qy   430 TNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVN---KP 485
          |||.||  || : || :| ||  | ||| ||||.| :.::..:|||||||..||   |
Db   441 TNIPFLLLLTNPVFIEGTYWTTFIDDTPQLFQMVSSQNRAQKLLHYLADVAVNGSSIKG 500

Qy   486 HGVRPKDVAAPIDKLPNIKDL--------PLPRGSRDRLKQLGPAAFARDLREQDALA 535
          . . ||: . .|.| |::.:|        .:|.|.   |::|||||||:.:|:|::.|
Db   501 QIGLPKLKSNP--SVPHLHDAQGNVINVTKSAPPSGWRQVLLEKGPAEFARQVRQFNGTL 558

Qy   536 VTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPW 595
          ::|||:|||||||||||||::. | .: ||.:|:.. .|:.||||||:||||||| ||||
Db   559 LMDTTWRDAHQSLLATRVRTHDLATIAPTTAHALAGRFALECWGGATFDVAMRFLHEDPW 618

Qy   596 DRLDELRERAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVS 655
          :||.:||  .|||:..||||||.|.| | .| :.: | .|:...::|||||||||||||.
Db   619 ERLRKLRSLVPNIPFQMLLRGANGVAYSSLPDNAIDHFVKQAKDNGVDIFRVFDALNDLE 678

Qy   656 QMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDM 715
          |::. :|||: .  .|  |. :.|||:.:|.:| |.|||||::|||||:.|:||| |||||
Db   679 QLKVGVDAVKKAG-GVVEATVCFSGDMLQPGKK-YNLDYYLEIAEKIVQMGTHILGIKDM 736

Qy   716 AGLLRPAAVTKLVTALRREF-DLPVHVTHTHDTAGGQLATYFAAAQAGADAVDGASAPLSG 774
          || ::|||. ||: :||.:| ||| |||||||:||.::|: .||| |||||||.|:..|.|
Db   737 AGTMKPAAAKLLIGSLRAKYPDLPIHVTHTHDSAGTAVASMTACALAGADVVDVAINSMSG 796

Qy   775 TTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYILPFESGTPGPTGRVYRHE 834
          .:|||| :|::.:| . :.| |:.: :.|:|.||.|:..| | ..|||.:..|..||.||
Db   797 LFSQPSINALLASL-EGNIDTGINVEHVRELDAYWAEMRLLYSCFEADLKGPDDPEVYQHE 855

Qy   835 IPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAG 894
          |||||| || .|| .|||.:.:. |.:||| ..:.|| :.||| ||||||||| ::|..
Db   856 IPGGQLTNLLFQAQOLGLGEQWAETKRAYREANYLLGDIVKVTPTSKVVGDLAQFMVSNK 915

Qy   895 VDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEE 954
          ..| |:  || ||||||:: .|: ||:|:| |||||| |:|:| :  . : || | :
Db   916 LTSDDVRRIANSLDFPDSVMDFFEGLIGQPYGGPPEPFRSDVLRNK---RRKLTCRPGLE 972

Qy   955 QAHLDADDSKERRNSLNRL----------LFPKPTEEFLEHRRRFGNTSALDDREFFY 1002
          : . | :.:.||::.|.||          ||| .|::|||:||||| |:|||||:|
Db   973 LEPFDLE--KIREDLQNRFGDVDECDVASYNMYPRVYEDFQKMRETYGDLSVLPTRSFLS 1030

Qy  1003 GLVEGRETLIRLPDVRTPLLVRLDAISEPDDK-GMRNVVANVNGQIRPMRVRDRSVESVT 1061
          .| : .|:.:.::..:. :|:|| |:.: :|| |:|::. :|.|::| :|||||.: .|
Db  1031 PLETDEEIEVVIEQGKT-LIIKLQAVGDLNKKTGEREVYFDLNGEMRKIRVADRSQKVET 1089

Qy  1062 ATAEKADSSNKGHVAAPFAGV-VTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKID 1120
          .|  |||.. . | :||:|||  .|.|.:.| ::||| .|.::::|||| |.:|.||::.
Db  1090 VTKSKADMHDPLHIGAPMAGVIVEVKVHKGSLIKKGQPVAVLSAMKMEMIISSPSDGQVK 1149

Qy  1121 RVVVPAATKVEGGDLIVV 1139
          .| :||.| :| |||:|:
Db  1150 EVFVSDGENVDSSDLLVLL 1168
```

```
RESULT 5
PYC2_YEAST
ID   PYC2_YEAST      STANDARD;      PRT;   1180 AA.
AC   P32327;
```

```
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyruvate carboxylase 2 (EC 6.4.1.1) (Pyruvic carboxylase 2) (PCB 2).
GN   PYC2 OR YBR218C OR YBR1507.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   STRAIN=S288c;
RX   MEDLINE=92017667; PubMed=1921979;
RA   Stucka R., Dequin S., Salmon J.-M., Gancedo C.;
RT   "DNA sequences in chromosomes II and VII code for pyruvate
RT   carboxylase isoenzymes in Saccharomyces cerevisiae: analysis of
RT   pyruvate carboxylase-deficient strains.";
RL   Mol. Gen. Genet. 229:307-315(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=96128067; PubMed=8554526;
RA   Dubois E., el Bakkoury M., Glansdorff N., Messenguy F., Pierard A.,
RA   Scherens B., Vierendeels F.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96128067; PubMed=8554526;
RA   Val D.L., Chapman-Smith A., Walker M.E., Cronan J.E. Jr.,
RA   Wallace J.C.;
RT   "Polymorphism of the yeast pyruvate carboxylase 2 gene and protein:
RT   effects on protein biotinylation.";
RL   Biochem. J. 312:817-825(1995).
```

```
DR   EMBL; X59890; CAA42544.1; -.
DR   EMBL; Z36087; CAA85182.1; -.
DR   EMBL; U35647; AAC49147.1; -.
DR   PIR; S46094; S46094.
DR   HSSP; P24182; 1BNC.
DR   SGD; S000422; PYC2.
DR   GO; GO:0005829; C:cytosol; IDA.
DR   GO; GO:0004736; F:pyruvate carboxylase activity; IDA.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPSase_L_D2.
DR   InterPro; IPR005481; CPSase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Ligase; Multifunctional enzyme; Biotin; Gluconeogenesis; ATP-binding;
KW   zinc; Multigene family.
FT   NP_BIND     183    188       ATP (POTENTIAL).
FT   ACT_SITE    313    313       BY SIMILARITY.
FT   BINDING    1136   1136       BIOTIN (BY SIMILARITY).
FT   CONFLICT     15     15       S -> C (IN REF. 1).
FT   CONFLICT    132    132       D -> E (IN REF. 1).
FT   CONFLICT    238    238       N -> K (IN REF. 1).
FT   CONFLICT    268    268       L -> F (IN REF. 1).
FT   CONFLICT    546    546       S -> C (IN REF. 1).
FT   CONFLICT    642    642       N -> T (IN REF. 1).
FT   CONFLICT    771    773       GTA -> STR (IN REF. 1).
FT   CONFLICT    831    831       W -> R (IN REF. 1).
FT   CONFLICT    839    839       Y -> N (IN REF. 1).
FT   CONFLICT   1001   1001       K -> R (IN REF. 1).
FT   CONFLICT   1155   1155       Q -> P (IN REF. 1).
FT   CONFLICT   1178   1178       K -> P (IN REF. 1).
FT   CONFLICT   1180   1180       K -> KVIFTR (IN REF. 1).
SQ   SEQUENCE   1180 AA;  130166 MW;  AD60DA3A60F5E001 CRC64;
```

```
Query Match             42.1%;  Score 2436;  DB 1;  Length 1180;
Best Local Similarity   45.5%;  Pred. No. 3.1e-123;
Matches  531;  Conservative 183;  Mismatches 408;  Indels 44;  Gaps 16;

Qy    7  STLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEG-   65
         |||||||||||||||  ||||  |  |||    |||  || ||||   | |  ||  |
Db   15  SLLGEKNKILVANRGEIPIRFRSAHELSMRTIAIYSHEDRLSMHRLKADEAYYIGEEGQ   74

Qy   66  -SPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLT  124
          ||||  ||||| ||| ||| ||||||||||||||   || ||||| ||| |||  ||
Db   75  YTPVGAYLAMDEIIEIAKKHKVDFIHPGYGFLSENSEFADKVVKAGITWIGPPAEVIDSV  134

Qy  125  GDKSRAVTAAKKAGLPVLAEST-PSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVAS  183
         |||  ||  ||| |||| ||| | ||  | ||||   ||| |||||| |||||||||| |
Db  135  GDKVSARHLAARANVPTVPGTPGPIETVQEALDFVNEYGYPVIIKAAFGGGGRGMRVVRE  194

Qy  184  PDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQLLGDHTGEVVHLYERDCSLQR  243
            |      ||   ||| |||   ||  |||   |||   |   ||  || ||| |  |
Db  195  GDDVADAFQRATSEARTAFGNGTCFVERFLDKPKHIEVQLLADNHGNVVHLFERDCSVQR  254

Qy  244  RHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRI  303
         ||||||| ||| ||||  |   || ||||  | ||| ||||||||   ||| |||||||
Db  255  RHQKVVEVAPAKTLPREVRDALLTDAVKLAKVCGYRNAGTAEFLVDNQNRHYFIEINPRI  314

Qy  304  QVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRP  363
         |||||| |   | ||||  |||  ||| ||  ||||||| ||||  |||||||| | | |
Db  315  QVEHTITEEITGIDIVSAQIQIAAGATLTQLGLLQDKITTRGFSIQCRITTEDPSKNFQP  374

Qy  364  DTGTITAYRSPGGAGVRLDGA-AQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAE  422
         |    | ||| |||| ||   | |    || | || ||| |       |    |  ||
Db  375  DTGRLEVYRSAGGNGVRLDGGNAYAGATISPHYDSMLVKCSCSGSTYEIVRRKMIRALIE  434

Qy  423  FTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTV  482
          |   |||| |||| | |   | ||| ||  |  | |  ||  |  ||    | ||| |
Db  435  FRIRGVKTNIPFLLTLLTNPVFIEGTYWTFIDDTPQLFQMVSSQNRAQKLLHYLADIAV   494

Qy  483  N----KPHGVRPKDVAAPIDKLPNIKDL------PLPRGSRDRLKOLGPAAFARDL     528
         |     |     ||    || |            |  | || |  | |   || |
Db  495  NGSSIKGQIGLPKLKSNP--SVPHLHDAQGNVINVYKSAPPSGWRQVLLEKGPSEFAKQV  552

Qy  529  REQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMR  588
          |||  |  ||| |||||||||||||  ||| || ||   ||  ||| | ||| || ||
Db  553  RQFNGTLLMDTTWRDAHQSLLATRVRTHDLATIAPTTAHALAGAFALECWGGATFDVAMR  612

Qy  589  FLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIF  648
         || | ||  |   ||| ||||| ||||| | | |     || | || | ||||||||||
Db  613  FLHEDPWERLRKLRSLVPNIPFQMLLRGANGVAYSSLPDNAIDHFVKQAKDNGVDIFRVF  672
```

```
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).

DR   EMBL; Y11106; CAA71993.1; -.
DR   HSSP; P24182; 1DV1.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005483; CPase_L.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   InterPro; IPR000891; HMGL-like.
DR   InterPro; IPR003379; PYC_OADA.
DR   InterPro; IPR005930; Pyruv_carbox.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPase_L_chain; 1.
DR   Pfam; PF02786; CPase_L_D2; 1.
DR   Pfam; PF00682; HMGL-like; 1.
DR   Pfam; PF02436; PYC_OADA; 1.
DR   PRINTS; PR00098; CPSASE.
DR   TIGRFAMs; TIGR01235; Pyruv_carbox; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Ligase; Multifunctional enzyme; Biotin; Gluconeogenesis; ATP-binding;
KW   zinc.
FT   NP_BIND     185    190    ATP (POTENTIAL).
FT   ACT_SITE    315    315    BY SIMILARITY.
FT   BINDING    1140   1140    BIOTIN (BY SIMILARITY).
SQ   SEQUENCE   1189 AA;   131400 MW;   8B6E85807965791A CRC64;

Query Match            41.3%;   Score 2391;   DB 1;   Length 1189;
Best Local Similarity  45.3%;   Pred. No. 8.2e-121;
Matches   526;   Conservative   177;   Mismatches   423;   Indels   34;   Gaps   16;

Qy      6  SSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEG   65
              |||| ||||||||||| : |:: :||| |||:   |||: :|||:|:  :  || ||| |
Db     16  SSLLGTMNKILVANRGEIPIRIFRTAHELSMNTVAIYSHEDRLSMHRLKADEAYVIGERG   75

Qy     66  --SPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDL  123
             |::|:|| :|||| : |.| ..|: :||||||:||:: ||:| |ENG| ::||:  :|||
Db     76  QYSPVQAYLAIDEIIKIAVKHNVNMIHPGYGFCSENSEFARKVEENGLIWVGPSDTVIDA  135

Qy    124  TGDKSRAVTAAKKAGLPVLAEST-PSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVA  182
            .|||.:|..  .||:  || |:| |.::::::|::|: : || |:|| |||||||||| |
Db    136  VGDKVSARNLAYAAANVPTVPGTPGPIEDVAQATAFVEEYGYPVIIKAAFGGGGRGMRVVR  195

Qy    183  SPDELRKLATEASREAEAAFGDGAVVVERAVINPQHIEVQILGDHTGEVHLYERDCSLQ   242
             .:: ||  ::||: ||: |||.| |.|| |::: :|||| | ||. |:|||||||| |
Db    196  EGDDIEDEAFLRASSEAKTAFGNGTVFIERFLDKPKHIEVQLLADNYGNVIHLFERDCSVQ  255

Qy    243  RRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPR  302
            |||||| : .  : | ||:|. | .|||:: :::.| ||||| ||||.:.|:: |||:||
Db    256  RRHQKVARNCSAKTLPEVRNAILNDAVKLAKTANYRNAGTAEFLVDSQNRHYFIEINPR   315

Qy    303  IQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFR  362
            |||||::||| | || ||||:: |||| |::|||.::||| :| | ||||||||| ::|:
Db    316  IQVEHTITEEITGVDIVAAQIQIAAGASLEQLGLLQEKITTRGFAIQCRITTEDPTKNFQ  375

Qy    363  PDTGITTAYRSPGGAGVRLDGAAQLGGE-ITAHFDSMLVKMTCRGSDFETAVARAQRALA  421
            ||||  . |||.||.|||||| .  | | |:||:|||||| : :||::.|||.| :|||
Db    376  PDTGKIEVYRSSGGNGVRLDGGNGFAGAVISPHYDSMLVKCSTSGSNYEIRRKMIRALV   435

Qy    422  EFTVSGVATNIGFLRALLREDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVT   481
            ||   ||:||| ||:|||:  . ::|||.||||||||||.|  || :||||||| .|:
Db    436  EFRIRGVKTNIPFLLALLTHPVFMTSECWTTFIDDTPELFKILTSQNRAQKLLAYLGDLA  495

Qy    482  VN--------------KPHGVRPKDVAAPIDKLPNIKD--LPLPRGSRDRLKQLGPAAFARDL  528
```

```
Qy    649  DALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAH  708
            |||| : |:  |:|| :    |  |  :  |||:|:| :|  |||||| :|: :| : |:
Db    673  DALNDLEQLKVGVNAVKKAG-GVVEATVCYSGDMLQPGKK-YNLDYYLEVVEKIVQMGTH  730

Qy    709  IIAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVHTHDTAGGQLATYFAAAQAGADAVDG  767
            ||.|||||| ::||| ::|:  :||::.||:|||:||:|:| : |:: .||:||||:||:
Db    731  ILGIKDMAGTMKPAAAKLLIGSLRTRYPDLPIHVHSHDSAGTAVASMTACALAGADVVDV  790

Qy    768  ASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPT  827
            | :.:SG:TSQPS:.|::|::  |: ||| .:| | : | :| | :: |:  |||   ||
Db    791  AINSMSGLTSQPSINALLASL-EGNIDTGINVEHVRELDAYWAEMRLLYSCFEADLKGPD  849

Qy    828  GRVYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVIPSSKVVGDLA  887
             .:|:HEIPGGQL:NL ||:  ||::::: ::|:|| |::    | ||||| |||||||||
Db    850  PEVYQHEIPGGQLTNLLFQAQQLGLGEQWAETKRAYREANYLLGDIVKVIPTSKVVGDLA  909

Qy    888  LHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPL  947
            :|:|.:.:.: ::: ::::  :||||: || |.|| |.|:|PEPLR::.  .: :..:|:
Db    910  QFMVSNKLTSDDIRRLANSLLDFPDSVMDFFEGLIGQPYGGFPEPLRSDVLRNK--RRKL  966

Qy    948  TEVPEEQAHLDADDSKERRNSLNRL--------LFPKPTEEFLEHRRRFGNTSAL      995
             .:| .     :.:  :||:: |||        |:||: ||:|:.| :||::.:|
Db    967  TCRPGLELEPFDLE--KIREDLQNRFGDIDECDVASYNMYPRVYEDFQKIRETYGDLSVL 1024

Qy    996  DDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDK-GMRNVVANVNGQIRPMRVRD 1054
             :::||:.   .:  ::||||:||||::|||:: | ||: |:| : |:|||:|: :||:|
Db   1025  PTKNFLAPAEPDEEIEVTIEQGKT-LIIKLQAVGDLNKKTGQREVYFELNGELRKIRVAD 1083

Qy   1055  RSVESVTATAFEKADDSSNKGHVAAPPAGV-VTVTVAEGDEVKAGDAVAIIEAMKMEAITA 1113
            :| :. .| ::  || : :  :..|: ||| :.: |: : :|:|| :|| ::||||| :::
Db   1084  KSQNIQSVAKPKADVHDTHQIGAPMAGVIIEVKVKHGSLVKKGESIAVLSAMKMEMVVSS 1143

Qy   1114  SVDGKIDRVVVPAATKVEGGDLIVVV 1139
             .|| ::. |.: || | ||.:|:||
Db   1144  PADGQVKDVFIKDGESVDASDLIVVL 1169
```

RESULT 6
PYC_PICPA

```
ID   PYC_PICPA     STANDARD;      PRT;   1189 AA.
AC   P78992;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Pyruvate carboxylase (EC 6.4.1.1) (Pyruvic carboxylase) (PCB).
GN   PYC1.
OS   Pichia pastoris (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Pichia.
OX   NCBI_TaxID=4922;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98301182; PubMed=9639311;
RA   Menendez J., Delgado J., Gancedo C.;
RT   "Isolation of the Pichia pastoris PYC1 gene encoding pyruvate
RT   carboxylase and identification of a suppressor of the Pyc
RT   phenotype.";
RL   Yeast 14:647-654(1998).
CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
CC       CARBOXYL GROUP TO PYRUVATE IN THE SECOND (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP + Pyruvate + HCO(3)(-) = ADP + phosphate +
CC       oxaloacetate.
CC   -!- COFACTOR: BIOTIN AND ZINC.
CC   -!- PATHWAY: Gluconeogenesis.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC       AND CARBAMYL PHOSPHATE SYNTHETASES.
CC   --------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
Db          RT   jannaschii.";
            RL   Science 273:1058-1073(1996).
            RN   [2]
            RP   SEQUENCE OF 1-12, AND FUNCTION.
            RX   MEDLINE=21034791; PubMed=11195096;
            RA   Mukhopadhyay B., Patel V.J., Wolfe R.S.;
            RT   "A stable archaeal pyruvate carboxylase from the hyperthermophile
            RT   Methanococcus jannaschii.";
            RL   Arch. Microbiol. 174:406-414(2000).
            CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
            CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
            CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE CARBOXYL
            CC       GROUP TO PYRUVATE IN THE SECOND.
            CC   -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
            CC       oxaloacetate.
            CC   -!- COFACTOR: ATP, MAGNESIUM (OR MANGANESE OR COBALT), PYRUVATE AND
            CC       BICARBONATE.
            CC   -!- ENZYME REGULATION: INHIBITED BY MAGNESIUM, WHEN ITS CONCENTRATION
            CC       EXCEEDED THE ATP ONE, AND BY HIGH CONCENTRATION OF ATP AND ALPHA-
            CC       KETOGLUTARATE.
            CC   -!- PATHWAY: Gluconeogenesis.
            CC   -!- SUBUNIT: HETEROOCTAMER OF FOUR A AND FOUR B SUBUNITS.
            CC   -!- MASS SPECTROMETRY: MW=55500; METHOD=MALDI.
            CC   -!- MISCELLANEOUS: ITS OPTIMUM PH IS 8.5 AND THE OPTIMUM TEMPERATURE
            CC       IS 80-90 DEGREES CELSIUS.
            CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
            CC       AND CARBAMYL PHOSPHATE SYNTHETASES.
            CC   --------------------------------------------------------------------------
            CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
            CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
            CC   the European Bioinformatics Institute. There are no restrictions on its
            CC   use by non-profit institutions as long as its content is in no way
            CC   modified and this statement is not removed. Usage by and for commercial
            CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
            CC   or send an email to license@isb-sib.ch).
            CC   --------------------------------------------------------------------------
            DR   EMBL; U67563; AAB99232.1; -.
            DR   PIR; D64453; D64453.
            DR   HSSP; P24182; 1BNC.
            DR   TIGR; MJ1229; -.
            DR   InterPro; IPR004549; AccC.
            DR   InterPro; IPR005482; Biotin_carb_C.
            DR   InterPro; IPR005483; CPase_L.
            DR   InterPro; IPR005479; CPase_L_D2.
            DR   InterPro; IPR005481; CPase_L_N.
            DR   Pfam; PF02785; Biotin_carb_C; 1.
            DR   Pfam; PF00289; CPase_L_chain; 1.
            DR   Pfam; PF02786; CPase_L_D2; 1.
            DR   PRINTS; PR00098; CPSASE.
            DR   TIGRFAMs; TIGR00514; accC; 1.
            DR   PROSITE; PS00866; CPSASE1; 1.
            DR   PROSITE; PS00867; CPSASE 2; 1.
            KW   Ligase; Multifunctional enzyme; Gluconeogenesis; Magnesium; Pyruvate;
            KW   ATP-binding; Complete proteome.
            FT   NP_BIND    162    167       ATP (POTENTIAL).
            FT   ACT_SITE   291    291       POTENTIAL.
            SQ   SEQUENCE   501 AA;  55402 MW;  04D2E401892F872F CRC64;

            Query Match             18.6%;  Score 1079;  DB 1;  Length 501;
            Best Local Similarity   48.2%;  Pred. No. 7e-51;
            Matches 218;  Conservative 82;  Mismatches 146;  Indels 6;  Gaps 4;

Qy     496  VNGSSIKGQIGLPKLH--KEADIPSITDINGDVIDVSIPPPDGWRQFLLKGPEQFAQQV 553
                ::|||||||| | :      |:|:| | ||||| ||: |:::|| |:|
Db     529  REQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMR 588

Qy     554  RAFPGLMIMDTTWRDAHQSLLATRVRTHDLLNIAPATSYALHHAFALECWGGATFDVSMR 613
                :||||||||:|| ||:: | :   | ||: |:||:| |:|:|| | | |::||||| |::|
Db     554  (blank)

Qy     589  FLFEDPWDRLDELRREAMPNVNIQMLLRGRNTVGYTPYPPDSVCRAFVKEAASSGVDIFRIF 648
                | | ||:| ||: | | ||: :||| :| || | || || || |: |::|
Db     614  FLHEDPWQRLRKLRKAVPNIPFSMLLRGGNGVAYYISLPDNAIDHFLKQAKDTGVFVFRVF 673

Qy     649  DALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAH 708
                |||| | :: ::  ::  :||   | | :  |
Db     674  DALNDIEQLKVGVDAVKKAG--GVVEATMCYSGDMLKPGKK-YNLEYYINLATEIVEMGTH 731

Qy     709  ILAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVTHDTAGGQLATYFAAAQAGADAVDG 767
                ||| | | |||| | ::|||::|: ::: |:|:: |:|::|: | :|||| | |: ::|||
Db     732  ILAVKDMAGTLKPTAAKQLISALRRKFPSLPIHVTHDSAGTGVASMVACARAGADVVTV 791

Qy     768  ASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLYLPFESGTPGPT 827
                ||:| |||:| | |::| | :  ||:||: :: :| |:| |::|| | :|| ||||:||
Db     792  RVNSMSGMTSQPSMSAFIASL-DGEIETGIPEANAREIDAYWAEMRLLYSCFEADLKGPD 850

Qy     828  GRVYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTIPSSKVVGDLA 887
                | : | ||||||:|: ||| | |:|   | || | | | ::|: | || ||||||||||
Db     851  PEVYQHEIPGGQLTNLLFQAQQVGLGEKWVETKKAYEAANRLLGDIVKVTPTSKVVGDLA 910

Qy     888  LHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSE--GK 944
                   |  | : : || ||:|  | |:| ||| | |||  ||||||:||::| : : : ||
Db     911  QFMVSNKLSSEDVERLASELDFPDSVLDFFEGLMGTPYCGFPEPLRTNVISGKRRKLTSR 970

Qy     945  APLTEVPEEQA--HLDADDSKERRNSL-NRLLFPKPEEFLEHRRRFPGNTSALDDREF 1000
                || : | |:   | ::| | :: | ||| :|| | | : |:| | :: ||| |
Db     971  PGLTLEPYNIPAIREDLEARFSKVTENDVASYNMYPKVYEAYKKQQELYGDLSVLPTRNF 1030

Qy    1001  FY--GLVEGRETLIRLPDVRTPLLVRLDAISE-PDDKGMRNVVANVNGQIRPMVRDRSV 1057
                ||   |:| | | : |::  :|||||  | :   ||||| | | | :| || ||||:||
Db    1031  LSPPKIDERHVTIVTIETRKTLIIKCMAEGELSQSSGTREVYFELNGEMRKKVTVEDKNG 1090

Qy    1058  ESVTATAEPKADSSNKGHVAAPFAGVVT-VTVAEGDEVKAGDAVAIIEAMKMEATITASVD 1116
                |:  :|  ||:  |: |:|  | | |  :  |::|| |:|| |  |||:||:|:| || |
Db    1091  AVETITRPKADAHNPNEIGAPMAGVVVEVRVHENGEVVKGDPIAVLSAMKMEMVISSPVA 1150

Qy    1117  GKIDRVVPAATKVEGGDLI 1136
                | |:|||: | || |:||
Db    1151  GRIGQIAVKENDSVDASDLI 1170


RESULT 7
PYCA_METJA
ID   PYCA_METJA     STANDARD;     PRT;     501 AA.
AC   Q58626;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pyruvate carboxylase subunit A (EC 6.4.1.1) (Pyruvic carboxylase A).
GN   PYCA OR MJ1229.
OS   Methanococcus jannaschii.
OC   Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC   Methanocaldococcaceae; Methanocaldococcus.
OX   NCBI_TaxID=2190;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX   MEDLINE=96337999; PubMed=8688087;
RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA   Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
```

```
Qy  191 ATEASREAFAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE 250
            |::|||||   ::|| :| ||||:|::||:||||:||:|||||:::|||::|||||||
Db  181 IESARNIAKSAFGDFTVFIEKYLENPRHIEQLLGDKHGNIIHLGDRECSIQRRHQKLIE 240

Qy  251 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 310
        :|| :   :|||  :||||::||: :|| ||||||||    ||:|::|:||:||||||||
Db  241 EAPSPIMTEELRERMGEAAIKAGAKAINYDSAGTVEFLY-ENGNFYFLEMNTRIQVEHTVT 299

Qy  311 EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 370
        |:||::||||| ::|:|||   |||:|::|::|:|||:||||||:|||::||:| |:|||
Db  300 EQVTGIDLVKAMIKIAAG---EELTLKQEDVKIRGHAIECRINAEDPLNDFVPCPGKIKL 356

Qy  371 YRSPGGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 430
        |||||  ||:|||  :||||  :::||::|  : |:|::||:|:|||||||:|||:|:||
Db  357 YRSPGPGVRIDSGYYGGAEIPPYYDSMIAKLIIYGNSREEAIARMKRALREYVIIGVKT 416

Qy  431 NIGFLRALLREEDFTSKRIATGFIADHPHLLQ 462
        ||||| ::||||:||  ||:| |::| :::|
Db  417 NIPFHRAVLEEENFLKGNISTHYVEQNMHKLR 448
```

RESULT 8
ACCC_ANASP
ID   ACCC_ANASP   STANDARD;   PRT;   447 AA.
AC   Q06862;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA
DE   carboxylase) (EC 6.4.1.2) (ACC).
GN   ACCC OR ALR0939.
OS   Anabaena sp. (strain PCC 7120).
OC   Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX   NCBI_TaxID=103690;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93352435; PubMed=8102363;
RA   Gornicki P., Scappino L.A., Haselkorn R.;
RT   "Genes for two subunits of acetyl coenzyme A carboxylase of Anabaena
RT   sp. strain PCC 7120: biotin carboxylase and biotin carboxyl carrier
RT   protein.";
RL   J. Bacteriol. 175:5268-5272(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21595285; PubMed=11759840;
RA   Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA   Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA   Yasuda M., Tabata S.;
RT   "Complete genomic sequence of the filamentous nitrogen-fixing
RT   cyanobacterium Anabaena sp. strain PCC 7120.";
RL   DNA Res. 8:205-213(2001).
CC   -!- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
CC   CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE
CC   CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE
CC   TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA.
CC   -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC   = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC   -!- PATHWAY: Long-chain fatty acid biosynthesis; first step.
CC   -!- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN
CC   CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
CC   OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX.
CC   -!- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
CC   PHOSPHATE SYNTHETASES.
CC   ---

DR   EMBL; L14862; AAB51770.1; -.
DR   EMBL; AP003584; BAB72896.1; -.
DR   PIR; A53311; A53311.
DR   PIR; AH1923; AH1923.
DR   HSSP; P24182; 1BNC.
DR   InterPro; IPR004549; AccC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   TIGRFAMs; TIGR00514; accC; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Fatty acid biosynthesis; Ligase; Biotin; ATP-binding;
KW   Complete proteome.
FT   NP_BIND   163   168   ATP (BY SIMILARITY).
FT   ACT_SITE   293   BY SIMILARITY.
SQ   SEQUENCE   447 AA; 49104 MW; 8A541B38B39E00F9 CRC64;

```
Query Match             17.9%;
Best Local Similarity   48.4%;   Pred. No. 1.1e-48;
Matches   216;   Conservative   73;   Mismatches   152;   Indels   5;   Gaps   3;

Qy   12 FKKILVANRGEIAVRAFRAALFTGAATVAITYPREDRGSFHRSFASEAVRIGTEGSPVKAY 71
         | ||:|||||||| | :|:|   |:|||:|     || ::|| | |::| |||| :|  ||
Db    3 FDKILIANRGEIALRILRACEMGIATIAVHSTVDRNALHVQLADEAVCIG-EPASAKSY 61

Qy   72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 131
         |:| :|| ||  ::|:|:||||||||||:::| ||| :  :|||||||:::|:||||:|:
Db   62 LNIPNIIAALTRNASAIHPGYGFLSENAKFAEICADHHIAFIGPTPEAIRLMGDKSTAK 121

Qy  132 TAAKKAGLPVLAESTPSKNID-EIVKSAEGQTYPIFVKAVAGGGRGMRFVASPDELRKL 190
         ::  |||:|   :: ::  :: | :  :: | |:::|||||||||||:|:|||||::||
Db  122 ETMQKAGVPTVPGSEGLVETEQEGLELAKDIGYPVMKATAGGGGRGMRLVRSPDEFVKL 181

Qy  191 ATEASREAFAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE 250
         |::|| ||||||: |||:|: ||:||:||||:||||:::|:|||||||||:|||:|::|
Db  182 FLAAQGEAGAAFGNAGVYIEKFIERPRHIEFQILADNYGNVIHLGERDCSIQRRNQLLE 241

Qy  251 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 310
         ||| : | | ||::|:|||||||:|:|| ||||:||| |:|:::|:|:||:||||||||
Db  242 EAPSPALDSDLREKMGQAAVKAAQFINYTGAGTIEFLLDRSGQFYFMEMNTRIQVEHPVT 301

Qy  311 EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 370
         | |||:|||:||:|:| |  |  ||:||:|:: :|||:||||| |||::|:|| | ||||
Db  302 EMVTGVDLLVEQIRIAQGERLR--LTQDQVVLRGHAIECRINAEDPDHDFRPAPGRISG 358

Qy  371 YRSPGGAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 430
         | |||  |||:| :  | | :::: | :  :  :|:|:|::: :  ||| |  :::| |
Db  359 YLPGGPGVRIDSHYTDYQIPPYDSLIGKLIVWGPDRATAINMKRALRECAITGLPT 418

Qy  431 NIGFLRALLREEDFTSKRIATGFIAD 456
         ||||::|::|||:|||||||||| ::
Db  419 TIGFHQRIMENPQFLQGNVSTSFVQE 444
```

RESULT 9
PYCA_ARCFU
ID   PYCA_ARCFU   STANDARD;   PRT;   506 AA.
AC   O30019;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pyruvate carboxylase subunit A (EC 6.4.1.1) (Pyruvic carboxylase A).
GN   PYCA OR AF0220.
OS   Archaeoglobus fulgidus.
OC   Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC   Archaeoglobaceae; Archaeoglobus.
OX   NCBI_TaxID=2234;

```
Db  121 ESMQRAGVPVIPGSPKIDTVDEAKEWAEKIGYPVAVKASGGGGIGIVVVNSQELEEAF 180
           :::|:||:||:  :||  |||     |||  |  ||| :|||  | :|| ::||
Qy  192 TEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVEI 251
           :: ::|:  |:|    :|| || |  |||  |||   | :|:| || : ||||||:
Db  181 RKSKKLGESYFKDSTVYLEKYLARPRHIEVQILADQHGNVIHLGERECSIQRRHQKLIEE 240
Qy  252 APAQHLDPELRDRICADAVKFCRSIGYQQAGTVEFIVDEKGNHVFIEMNPRIQVEHTVTE 311
          ||: :|:  | : |  :|:|  |:||| :|||:||:| ||| ::|:|| |||||||:||
Db  241 APSPALNEMREELGKLAVKGAREIGYTNAGTEFFLY-ENGNFYFLEINSRLQVEHTITE 299
Qy  312 EVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITEDPNNGFRPDTGTTAY 371
          :||  |:|| |:|:| |: :  |    ::: |  |  :|| ||||:|  |:| :||:|
Db  300 VVTGIDIVKYQIRIAYG---EELRHGQEDVAIRGHAIECRINAEDPVN-FYPRSGRILHY 355
Qy  372 RSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATN 431
          ||||:|:|:| |  :||| |  | |||:|| | || | |:||: |:||| |: :|:| ||
Db  356 RSPGGIGIRVDSGIHMGYRIPEEYDSMISKLIAYGETREEAIARMKRALYEYIIEGVETN 415
Qy  432 IGFLRALLREEDFTSKRIATGFIAD 456
          |:|: |::::|:| ::  : ||::
Db  416 IPFHFAVLNDEEFVRGNIHTKFVEE 440
```

RESULT 10
PYCA_METTH
```
ID   PYCA_METTH   STANDARD;      PRT;    491 AA.
AC   O27939;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pyruvate carboxylase subunit A (EC 6.4.1.1) (Pyruvic carboxylase A).
GN   PYCA OR MTH1917.
OS   Methanobacterium thermoautotrophicum.
OC   Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC   Methanobacteriaceae; Methanothermobacter.
OX   NCBI_TaxID=187420;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Delta H;
RX   MEDLINE=98037514; PubMed=9371463;
RA   Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA   Aldredge T.R., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA   Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA   Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA   Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA   McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA   Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT   "Complete genome sequence of Methanobacterium thermoautotrophicum
RT   deltaH: functional analysis and comparative genomics.";
RL   J. Bacteriol. 179:7135-7155(1997).
RN   [2]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-17.
RC   STRAIN=Delta H;
RX   MEDLINE=98148063; PubMed=9478969;
RA   Mukhopadhyay B., Stoddard S.F., Wolfe R.S.;
RT   "Purification, regulation, and molecular and biochemical
RT   characterization of pyruvate carboxylase from Methanobacterium
RT   thermoautotrophicum strain deltaH.";
RL   J. Biol. Chem. 273:5155-5166(1998).
CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE CARBOXYL
CC       GROUP TO PYRUVATE IN THE SECOND. THE MAXIMUM ACTIVITY IS AT PH 8
CC       AND 60 DEGREES CELSIUS.
CC   -!- CATALYTIC ACTIVITY: ATP + Pyruvate + HCO(3)(-) = ADP + phosphate +
CC       oxaloacetate.
CC   -!- COFACTOR: ATP, MAGNESIUM (OR MANGANESE OR COBALT), PYRUVATE AND
CC       BICARBONATE.
CC   -!- ENZYME REGULATION: INHIBITED BY ADP AND ALPHA-KETOGLUTARATE.
CC   -!- PATHWAY: Gluconeogenesis.
CC   -!- SUBUNIT: HETEROOCTAMER OF FOUR A AND FOUR B SUBUNITS.
CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX   MEDLINE=98049343; PubMed=9389475;
RA   Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA   Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA   Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA   Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA   Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA   Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA   Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA   Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA   Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA   Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA   Venter J.C.;
RT   "The complete genome sequence of the hyperthermophilic, sulphate-
RT   reducing archaeon Archaeoglobus fulgidus.";
RL   Nature 390:364-370(1997).
CC   -!- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC       INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC       ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE CARBOXYL
CC       GROUP TO PYRUVATE IN THE SECOND.
CC   -!- CATALYTIC ACTIVITY: ATP + Pyruvate + HCO(3)(-) = ADP + phosphate +
CC       oxaloacetate.
CC   -!- COFACTOR: ATP, MAGNESIUM (OR MANGANESE OR COBALT), PYRUVATE AND
CC       BICARBONATE (BY SIMILARITY).
CC   -!- PATHWAY: Gluconeogenesis.
CC   -!- SUBUNIT: HETEROOCTAMER OF FOUR A AND FOUR B SUBUNITS (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC       AND CARBAMYL PHOSPHATE SYNTHETASES.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; AE001090; AAB91012.1; -.
DR   PIR; D69277; D69277.
DR   HSSP; P24182; 1BNC.
DR   TIGR; AF0220; -.
DR   InterPro; IPR004549; AccC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   TIGRFAMs; TIGR00514; accC; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; FALSE_NEG.
KW   Ligase; Multifunctional enzyme; Gluconeogenesis; Magnesium; Pyruvate;
KW   ATP-binding; Complete proteome.
FT   NP_BIND     161    166    ATP (POTENTIAL).
FT   ACT_SITE    290    290    POTENTIAL.
SQ   SEQUENCE   506 AA;  57438 MW;  39B765F319235AD1 CRC64;
```

```
Query Match        17.6%;  Score 1019;  DB 1;  Length 506;
Best Local Similarity   47.9%;  Pred. No. 1.2e-47;
Matches  213;  Conservative  73;  Mismatches 153;  Indels  6;  Gaps  4;

Qy   12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 71
         | |||||||||||| ||  :||||  |||  | ::||||: ||: ||||: :|:
Db    2 FSKILVANRGEIAVRVMRACRELGIKTVGVYSSADKRAFHRVVADECYIG-KADPRDSY 60
Qy   72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 131
         |:|:||  ||| |||:|:||||||||||| :|||:| ||:|||| |||: : :||:|
Db   61 LNIDRIIEVAKKSGAEAIHPGYGFLAENAEFAERCEEEGIVFIGPSPEVIRIAGSKVRSR 120
Qy  132 TAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKLA 191
```

```
CC      AND CARBAMYL PHOSPHATE SYNTHETASES.
CC      -------------------------------------------------------------
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
DR      EMBL; AE000942; AAB86377.1; -.
DR      PIR; A69123; A69123.
DR      HSSP; P24182; 1BNC.
DR      InterPro; IPR004549; AccC.
DR      InterPro; IPR005482; Biotin_carb_C.
DR      InterPro; IPR005479; CPase_L_D2.
DR      InterPro; IPR005481; CPase_L_N.
DR      Pfam; PF02785; Biotin_carb_C; 1.
DR      Pfam; PF00289; CPSase_L_chain; 1.
DR      Pfam; PF02786; CPSase_L_D2; 1.
DR      TIGRFAMs; TIGR00514; accC; 1.
DR      PROSITE; PS00866; CPSASE_1; 1.
DR      PROSITE; PS00867; CPSASE_2; 1.
KW      Ligase; Multifunctional enzyme; Gluconeogenesis; Magnesium; Pyruvate;
KW      ATP-binding; Complete proteome.
FT      NP_BIND      162    167       ATP (BY SIMILARITY).
FT      ACT_SITE     291    291       POTENTIAL.
SQ      SEQUENCE   491 AA;  54656 MW;  5789C34DA475C2E CRC64;

Query Match             17.2%;  Score 997;  DB 1;  Length 491;
Best Local Similarity   45.9%;  Pred. No. 1.7e-46;
Matches  205;  Conservative   79;  Mismatches  157;  Indels    6;  Gaps    4;

Qy       12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKRAY 71
             ||||||||||||||||||||||||||||||||| ||    :||  :||  ||| :| ::|
Db        2 FSKILVANRGEIAIRVMRACRELGIKSVAVYSEADKNALFTRYADEAYFIG-KPAPSQSY 60

Qy       72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 131
             | |||  :||   :||||||||||||| || :  ::||    | ::||||||| ::| :
Db       61 LRIDRILEVAEKAGAEAIHPGYGFLAENPRLGEECEKQGIKLIGPKGSVIEAMGDKITSK 120

Qy      132 TAAKKAGLPVLAESTPS-KNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKL 190
             ||||   ||:||:   :  |:  ||  : ::|  :||||  |||| ::| |
Db      121 KLMKKAGVPVIPGTDGGVSDPDEAARIADSIGYPVIIKASAGGGGIGMRAVYEEDELIRA 180

Qy      191 ATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE 250
             || :|  |  ||    | :  ||:  :  ||| :|::  ||| | :  : |: :|  ::|
Db      181 MESTQSVAASAFGDPTVYIEKYLERPRHIEFQVMADESGNVIHLADRECSIQRRHQKLIE 240

Qy      251 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 310
             ||: : |||||:::  ||| |||||||| ||||  ||||||||||| :||||||||| :|
Db       241 EAPSPIMTPELRERMGSAAVKAAEYIGYENAGTVEFLY-SNGDFYFLEMNTRIQVEHPIT 299

Qy      311 EEVTEVDLVKAQMRLAAGATLKELGLITQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 370
              || : ||||  :|  ||   ||  :    | ||||| ||  | ||   |  |  ||
Db       300 EVITGVDLVKEQIRVASG---EELRFTQKDINIRGHATECRINAENPLADFAPNPGKITG 356

Qy      371 YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 430
             |||||| | || || :|| | ||||:|:||| |  :| ::| :||  :::| | ::||:|
Db       357 YRSPGGIGVRVDSGVYMNYEIPPFYDSMISKLIVWGMDRQEAINRMKKRALSEYIILGVKT 416

Qy      431 NIGFLRALLREEDFTSKRIATGFIADH 457
             ||| ||| ::  : ::|::  : ::::
Db       417 TIPFHKAIMRNEAFRRGELHTHFVDEY 443

RESULT 11
ACCC_HAEIN
ID   ACCC_HAEIN     STANDARD;      PRT;   448 AA.
AC   P43873;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
```

```
DE   Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA
DE   carboxylase) (EC 6.4.1.2) (ACC).
GN   ACCC OR HI0972.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Rd / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT   Rd.";
RL   Science 269:496-512(1995).
CC   -!- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
CC       CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE
CC       CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE
CC       TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC       = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC   -!- PATHWAY: Long-chain fatty acid biosynthesis; first step.
CC   -!- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN
CC       CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
CC       OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
CC       PHOSPHATE SYNTHETASES.
CC   -------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; U32778; AAC22632.1; -.
DR   PIR; F64105; F64105.
DR   HSSP; P24182; 1BNC.
DR   TIGR; HI0972; -.
DR   InterPro; IPR004549; AccC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   TIGRFAMs; TIGR00514; accC; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Fatty acid biosynthesis; Ligase; Biotin; ATP-binding;
KW   Complete proteome.
FT   NP_BIND      163    168       ATP (POTENTIAL).
FT   ACT_SITE     292    292       POTENTIAL.
SQ   SEQUENCE   448 AA;  49108 MW;  2B497E2A31ED96D1 CRC64;

Query Match             16.3%;  Score 946;  DB 1;  Length 448;
Best Local Similarity   47.6%;  Pred. No. 8.3e-44;
Matches  208;  Conservative   61;  Mismatches  154;  Indels   14;  Gaps    7;

Qy       13 KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL 72
             :||::|||||||||||||||||||||||| :||||:|||::||||||||||||||:|
Db        3 EKVVIANRGEIALRLRACKELGIKTVAVHSTADRDLKHVLLADETICIGPAPS-AKSYL 61

Qy       73 DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT 132
```

```
Db         :  ||  ||.:    |||| .:|||||||||| |  |
    62  NIPAIIAAAEVTGADAIIHPCYGFLSENADFAEQVERSGFTFIGPTADVIRLMGDKVSAIK 121

Qy        :||.:     |||||||:     ||  .:  ||||||  |  |.:  |  .|  :. |||| |
   133  AAKKAGLPVL--AESTPSKNIDEIVKSABGQTYPIFVKAVAGGGRGMRFVASPDELRKL 190
Db        ||||:: :  .: |||.:  .|  ||||||| | | |.:  ||
   122  AMKKAGVPCVPGSDGPVSNDIAKNKEIAKRIGYPIIIKASGGGGRGMRVVRSEDALEES 181

Qy  ATEASREAAAFGDGAVYYERAVINPQHIEVQILGDHTGEVHLYERDCSLQRRHQKVVE 250
Db        |||||||  |||||||||:|||||:||||||||||||||:||||:||:||:||||||||
   182  IAMTKAEAKAAFNNDMVYMEKYLENPRHVEIQVLADTHGNAVILAERDCSMQRRHQKVVE 241

Qy  TAPAQHLDPELRDRI---CADAVKFCRSIGYQQGAGTVEFLVDEKGNHVFIEMNRIQVEH 307
Db        ||| .|  |.:    |:|||  |  ||||| |  ||||| ||||||:|||:|||
   242  EAPAPGITEVRRDIGSRCANA---CVEIGYRGAGTFEFLY-ENGEFYFIEMNTRIQVEH 297

Qy  TVIEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGT 367
Db        |:  ||   |||:  ||  |.:   |||   |.||  |  |.: |||   ||
   298  PVTEMITGVDLVKEQLRIAAGL---PISFKQEDIKVKGHAMECRINAEDPKT-FLPSPGK 353

Qy  ITAYRSPGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSG 427
Db        :||   ||||| |||| | |  ||||:||  |.:  ||.:|  ||  |  ||:||  |.:
   354  VNHLHSPGGLGVRWDSHVYGGYTVPPHYDSMIAKLITYGDTREVAIRRMQNALSETIIDG 413

Qy  VATNIGFLRALLREDF 444
Db        | ||  |:|  |.: :|:|
   414  IKNIPLHELILEDENF 430
```

RESULT 12

ACCC_ECOLI
ID   ACCC_ECOLI   STANDARD;   PRT;   449 AA.
AC   P24182;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA
DE   carboxylase) (EC 6.4.1.2) (ACC).
GN   ACCC OR FABG OR B3256.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=K12;
RX   MEDLINE=92112819; PubMed=1370469;
RA   Li S.-J., Cronan J.E. Jr.;
RT   "The gene encoding the biotin carboxylase subunit of Escherichia coli
RT   acetyl-CoA carboxylase.";
RL   J. Biol. Chem. 267:855-863(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92052166; PubMed=1682920;
RA   Kondo H., Shiratsuchi K., Yoshimoto T., Masuda T., Kitazono A.,
RA   Tsuru D., Anai M., Sekiguchi M., Tanabe T.;
RT   "Acetyl-CoA carboxylase from Escherichia coli: gene organization and
RT   nucleotide sequence of the biotin carboxylase subunit.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:9730-9733(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Best E.A., Knauf V.C.;
RL   Submitted (XXX-1992) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [5]
RP   SEQUENCE OF 1-12.
RC   STRAIN=K12 / EMG2;
RX   MEDLINE=97443975; PubMed=9298646;
RA   Link A.J., Robison K., Church G.M.;
RT   "Comparing the predicted and observed properties of proteins encoded
RT   in the genome of Escherichia coli K-12.";
RL   Electrophoresis 18:1259-1313(1997).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX   MEDLINE=94347758; PubMed=7915138;
RA   Waldrop G.L., Rayment I., Holden H.M.;
RT   "Three-dimensional structure of the biotin carboxylase subunit of
RT   acetyl-CoA carboxylase.";
RL   Biochemistry 33:10249-10256(1994).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX   MEDLINE=20283656; PubMed=10821865;
RA   Thoden J.B., Blanchard C.Z., Holden H.M., Waldrop G.L.;
RT   "Movement of the biotin carboxylase B-domain as a result of ATP
RT   binding.";
RL   J. Biol. Chem. 275:16183-16190(2000).
CC   -!- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
CC   CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE
CC   CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE
CC   TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA.
CC   -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC   = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC   -!- PATHWAY: Long-chain fatty acid biosynthesis; first step.
CC   -!- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN
CC   CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
CC   OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX.
CC   -!- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
CC   PHOSPHATE SYNTHETASES.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; M79446; AAA23748.1; -.
DR   EMBL; M80458; AAA23409.1; -.
DR   EMBL; M83198; AAA23746.1; -.
DR   EMBL; U18997; AAA58059.1; -.
DR   EMBL; AE000404; AAC76288.1; -.
DR   PIR; JS0632; JS0632.
DR   PDB; 1BNC; 30-AUG-95.
DR   PDB; 1DV1; 09-JUN-00.
DR   PDB; 1DV2; 09-JUN-00.
DR   PDB; 1K69; 05-DEC-01.
DR   EcoGene; EG10276; accC.
DR   InterPro; IPR004549; AccC.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR005479; CPase_I_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   TIGRFAMs; TIGR00514; accC; 1.
DR   PROSITE; PS00866; CPSASE_1; 1.
DR   PROSITE; PS00867; CPSASE_2; 1.
KW   Fatty acid biosynthesis; Ligase; Biotin; ATP-binding; 3D-structure;
KW   Complete proteome.
FT   NP_BIND   163   168   ATP (POTENTIAL).
FT   ACT_SITE   292   292   POTENTIAL.
FT   CONFLICT   260   261   CA -> SR (IN REF. 2).
FT   CONFLICT   313   313   L -> M (IN REF. 1).
FT   STRAND   4   7
FT   HELIX   11   24
FT   TURN   25   25
FT   STRAND   27   33

```
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC    the European Bioinformatics Institute.  There are no restrictions on its
CC    use by  non-profit  institutions as long  as its content  is in no  way
CC    modified and this statement is not removed.  Usage  by  and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    ----------------------------------------------------------------------------
GN    ACCC OR FABG OR PA4848.
OS    Pseudomonas aeruginosa.
OC    Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC    Pseudomonadaceae; Pseudomonas.
OX    NCBI_TaxID=287;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=ATCC 15692 / PAO1;
RX    MEDLINE=94042851; PubMed=7693652;
RA    Best E.A., Knauf V.C.;
RT    "Organization and nucleotide sequences of the genes encoding the
RT    biotin carboxyl carrier protein and biotin carboxylase protein of
RT    Pseudomonas aeruginosa acetyl coenzyme A carboxylase.";
RL    J. Bacteriol. 175:6881-6889(1993).
RN    [2]
RP    SEQUENCE FROM N.A.
RC    STRAIN=ATCC 15692 / PAO1;
RX    MEDLINE=20437337; PubMed=10984043;
RA    Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA    Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA    Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA    Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA    Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA    Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT    "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT    opportunistic pathogen.";
RL    Nature 406:959-964(2000).
CC    -!- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
CC        CARBOXYLASE COMPLEX: FIRST, BIOTIN CARBOXYLASE CATALIZES THE
CC        CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE
CC        TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA.
CC    -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC        = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC    -!- PATHWAY: Long-chain fatty acid biosynthesis; first step.
CC    -!- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN
CC        CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
CC        OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX.
CC    -!- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
CC        PHOSPHATE SYNTHETASES.
CC    ----------------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    ----------------------------------------------------------------------------
DR    EMBL; L14612; AAA16041.1; -.
DR    EMBL; AE004898; AAG08233.1; -.
DR    PIR; B49342; B49342.
DR    HSSP; P24182; 1BNC.
DR    InterPro; IPR004549; AccC.
DR    InterPro; IPR005482; Biotin_carb_C.
DR    InterPro; IPR005479; CPase_L_D2.
DR    InterPro; IPR005481; CPase_L_N.
DR    Pfam; PF02785; Biotin_carb_C; 1.
DR    Pfam; PF00289; CPsase_L_chain; 1.
DR    Pfam; PF02786; CPSase_L_D2; 1.
DR    TIGRFAMs; TIGR00514; accC; 1.
DR    PROSITE; PS00866; CPSASE_1; 1.
DR    PROSITE; PS00867; CPSASE_2; 1.
KW    Fatty acid biosynthesis; Ligase; Biotin; ATP-binding;
KW    Complete proteome.
FT    NP_BIND      163    168       ATP (POTENTIAL).
FT    ACT_SITE     292    292       POTENTIAL.
SQ    SEQUENCE   449 AA;  48888 MW;  3B04C77785C73541 CRC64;
```

```
Query Match             Best Local Similarity   16.0%;   Score 924;  DB 1;  Length 449;
                        Conservative 45.9%;  Pred. No. 1.3e-42;
Matches 199;  Conservative 67;  Mismatches 160;  Indels 8;  Gaps 5;

Qy           13 KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL 72
```

```
Query Match             Best Local Similarity   16.2%;   Score 936;  DB 1;  Length 449;
                        Conservative 46.2%;  Pred. No. 2.9e-43;
Matches 200;  Conservative 69;  Mismatches 156;  Indels 8;  Gaps 5;

Qy           14 KILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLD 73
                ||:||||||||:||:: |:|||:: ||| ||||||||||| |:||||||||||||||||
Db            4 KIVIANRGEIALRILRACKELGIKTVAVHSSADRDLKHVLLADETVCIGPAPS-VKSYLN 62

Qy           74 IDEIIGAAKKVKADAIYPGVGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVTA 133
                | ||||  | : |||:: | |||||| |||| | ||||||| ||||||||:||| ||||
Db           63 IPAIISAAEITGAVAIHPGVGFLSENANFAEQVERSGFIFIGPKAETIRLMGDKVSAIAA 122

Qy          134 AKKAGLPVL--AESTPSKNIDEIVKSAEGCTYPIFVKAVAGGGRGMRFVASPDELRKLA 191
                : |||  |:  |:: :|| :| || :|  ||:: :|| |||   ||   | :|| || |
Db          123 MKKAGVPCVPVGSDGPLGDDMDKNRAIAKRIGYPVIIKASGGGGRGMRVVRGDAELAQSI 182

Qy          192 TEASREAAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVEI 251
                : |||:||| ||: | ||   |  |::|||::|| |||||| ||| |||| || ||| |
Db          183 SMTRAEAKAAFSNDMVYMEKYLENPRHVEIQVLADGQGNSIYLAERDCSMQRRHQKVVEE 242

Qy          252 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE 311
                || | | || | :|| || |: |   ||| | ||| |||  |||| ||| |||| |: |
Db          243 APAPGITPELRRYIGERCAKACVDIGYRGAGTTEFLF-ENGEFYFIEMNTRIQVEHPVTE 301

Qy          312 EVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAY 371
                || | | : :| |:|| || ||| :|:  ||  | | :||::::||:  ||:: : ||:|
Db          302 MIITGVDLIKEQLRIAAG---QPLSIKQEEVHVRGHAVECRINAEDPNT-FLPSPGKITRF 357

Qy          372 RSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATN 431
                |   | | |:||:|:|||: |||:  : :| :| |  :|:|:|| || |||   ||| :|
Db          358 HAPGGFGVRWESHIYAGYTVPPYYDSMICKLICYGENRDVAIARMKNALQELIIDGIKTN 417

Qy          432 IGFLRALLREDF 444
                : :|: |:|::|
Db          418 VDLQIRIMNDENF 430
```

```
RESULT 14
ACCC_PSEAE
ID    ACCC_PSEAE     STANDARD;      PRT;   449 AA.
AC    P37798;
DT    01-OCT-1994 (Rel. 30, Created)
DT    01-OCT-1994 (Rel. 30, Last sequence update)
DT    28-FEB-2003 (Rel. 41, Last annotation update)
DE    Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA
DE    carboxylase) (EC 6.4.1.2) (ACC).
```

```
Db         :|:|!|||||||:| ||     | |||:  ||||      :|:||
        3  EKVLIANRGEIALRILRACKELGIKTVAVHSTADRELMHLSLADESVCIG-PAPATQSYL   61

Qy
Db        73  DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT  132
           ||:   |:  :|:||||||:| ||||||:|:|||||||  ||| |:|||
Db        62  QIPAIIAAAEVTGATAIHPGYGFLAENADFAEQIERSGFTFVGPTAEVIRLMGDKVSAKD  121

Qy       133  AAKKAGLPVLAESTPSKNIDE-IVKSAEGQTYPIFVKAVAGGGRGMRFVASPDELRKL   190
           |:|:|:|  || : |||||||  |:: |||||||||||||||||
Db       122  AMKRAGVPTVPGSDGPLPEDEETALAIAREVGYPVIIKAAGGGGRGMRVVYDESELIKS  181

Qy       191  ATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE  250
           ||||  || :||| | ||||||:|||||:| | :||:|||||||||||||||||:
Db       182  AKLTRTEAGAAFGNPMVYLEKFLTNPRHVEVQVLSDGQGNAIHLGDRDCSLQRRHQKVIE  241

Qy       251  IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT  310
           ||| |:: || :|||:|||  :| ||:  ||||||||||:|:| :|||||||:|||:|
Db       242  EAPAPGIDEKARQEVFARCVQACIEIGYRGAGTFEFLY-ENGRFYFIEMNTRVQVEHPVS  300

Qy       311  EEVITEVDLVKAQMRLAAGATLKELGLITQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA  370
           |:||:|| :| | :|||  :|| |:||||||||||||||:||||||||||||:||||:||:
Db       301  EMVTGVDIVKEMLRIASG---EKLSIRQEDVVIRGHALECRINAEDPKT-FMPSPGKVKH  356

Qy       371  YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT  430
           |: : | || || | :||:  |:|| ||: |: |  ||:  |  ||| |:: |||| |
Db       357  FHAPGGNGVRVDSHLYSGYSVPPNVDSLVGKVITYGADRDEALARMRNALDELIVDGIKT  416

Qy       431  NIGFLRALLREEDF  444
           | | |  |:|||:
Db       417  NTELHKDLVRDAAF  430
```

RESULT 15
MCCA_HUMAN
ID   MCCA_HUMAN   STANDARD;   PRT;   725 AA.
AC   Q96RQ3; Q9H959; Q9NS97;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Methylcrotonoyl-CoA carboxylase alpha chain, mitochondrial precursor
DE   (EC 6.4.1.4) (3-Methylcrotonyl-CoA carboxylase 1) (MCCase alpha
DE   subunit) (3-methylcrotonyl-CoA:carbon dioxide ligase alpha subunit).
GN   MCCC1 OR MCCA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND VARIANTS MCGI ARG-325 AND SER-385.
RX   PubMed=11170088;
RA   Gallardo M.E., Desviat L.R., Rodriguez J.M., Esparza-Gordillo J.,
RA   Perez-Cerda C., Perez B., Rodriguez-Pombo P., Criado O., Sanz R.,
RA   Morton D.H., Gibson K.M., Le T.P., Ribes A., Rodriguez de Cordoba S.,
RA   Ugarte M., Penalva M.A.;
RT   "The molecular basis of 3-methylcrotonylglycinuria, a disorder of
RT   leucine catabolism.";
RL   Am. J. Hum. Genet. 68:334-346(2001).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT MCGI HIS-464.
RX   MEDLINE=21295033; PubMed=11401427;
RA   Obata K., Fukuda T., Morishita R., Abe S., Asakawa S., Yamaguchi S.,
RA   Yoshino M., Ihara Y., Murayama K., Shigemoto K., Shimizu N., Kondo I.;
RT   "Human biotin-containing subunit of 3-methylcrotonyl-CoA carboxylase
RT   gene (MCCA): cDNA sequence, genomic organization, localization to
RT   chromosomal band 3q27, and expression.";
RL   Genomics 72:145-152(2001).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANT MCGI PHE-535.
RX   MEDLINE=21294419; PubMed=11406611;
RA   Holzinger A., Roeschinger W., Lagler F., Mayerhofer P.U., Lichtner P.,
RA   Kattenfeld T., Thuy L.P., Nyhan W.L., Koch H.G., Muntau A.C.,
RA   Roscher A.A.;
RT   "Cloning of the human MCCA and MCCB genes and mutations therein reveal

RT   the molecular cause of 3-methylcrotonyl-CoA: carboxylase
RT   deficiency.";
RL   Hum. Mol. Genet. 10:1299-1306(2001).
RN   [4]
RP   SEQUENCE FROM N.A., AND VARIANTS MCGI VAL-289; SER-385; PRO-437 AND
RP   HIS-532.
RX   PubMed=11181649;
RA   Baumgartner M.R., Almashanu S., Suormala T., Obie C., Cole R.N.,
RA   Packman S., Baumgartner E.R., Valle D.;
RT   "The molecular basis of human 3-methylcrotonyl-CoA carboxylase
RT   deficiency.";
RL   J. Clin. Invest. 107:495-504(2001).
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA   Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA   Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skeletal muscle;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- CATALYTIC ACTIVITY: ATP + 3-methylcrotonyl-CoA + HCO(3)(-) = ADP +
CC   phosphate + 3-methylglutaconyl-CoA.
CC   -!- COFACTOR: Biotin.
CC   -!- PATHWAY: Leucine catabolism.
CC   -!- SUBUNIT: Probably a dodecamer composed of six biotin-containing
CC   alpha subunits and six beta subunits.
CC   -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC   -!- DISEASE: Defects in MCCC1 are the cause of 3-
CC   methylcrotonylglycinuria type I (MCGI) [MIM:210200]; also
CC   designated CGA or CG2. MCGI is a recessive disease that is
CC   characterized by muscular hypotonia and atrophy, probably of
CC   spinal origin.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; AF310972; AAG53095.1; -.
DR   EMBL; AB029826; BAA99407.1; -.
DR   EMBL; AF297332; AAK67986.1; -.
DR   EMBL; AF310339; AAG50245.1; -.
DR   EMBL; AK023051; BAB14377.1; -.
DR   EMBL; BC004214; AAH04214.1; -.
DR   EMBL; BC004187; AAH04187.1; -.
DR   Genew; HGNC:6936; MCCC1.

```
Qy  363 PDTGTITAYRSP-GGAGVRLDGAAQLGGEITAHFDSMLYKMTCRGSDFETAVARAQRALA 421
            |: |  :  || :  |  :||| |  : :   |: |: ::|||| | : : :|
Db  397 PVAGPLVHLSTPRADPSTRIETGVRQGDEVSVHYDPMIAKLVVWAADRQAALTKLRYSLR 456

Qy  422 EFTVSGVATNIGFLRALLREEDFTISKRIATGFIADHPHLLQAPPADDEQGRILDYLADVT 481
            :: :: ||| ||| |||  :  : :|  | |
Db  457 QYNIVGLPTNIDFLLNLSGHPEFEAGNVHTDFIPQH---------------------- 492

Qy  482 VNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTTF 541
            ---H-----------KQLLLSRKAAAK-ESLCQAALGLILKEK--AMTDTFT 527
Db  493

Qy  542 RDAH--------QSLLATRVRSFALKPAAEAVA 566
            :: :|: :  : : | :|  ||
Db  528 LQAHDQFSPFSSSSGRRLNISYTRNMTLKDGKNNVA 563
```

Search completed: September 24, 2003, 15:56:06
Job time : 21.3631 secs

```
DR   GK; Q96RQ3; -.
DR   MIM; 210200; -.
DR   GO; GO:0005759; C:mitochondrial matrix; NAS.
DR   GO; GO:0009374; F:biotin binding activity; NAS.
DR   GO; GO:0004485; F:methylcrotonyl-CoA carboxylase activity; NAS.
DR   GO; GO:0006768; P:biotin metabolism; NAS.
DR   GO; GO:0006552; P:leucine catabolism; NAS.
DR   InterPro; IPR001882; Biotin_attach.
DR   InterPro; IPR005482; Biotin_carb_C.
DR   InterPro; IPR000089; Biotin_lipoyl.
DR   InterPro; IPR005479; CPase_L_D2.
DR   InterPro; IPR005481; CPase_L_N.
DR   Pfam; PF02785; Biotin_carb_C; 1.
DR   Pfam; PF00364; biotin_lipoyl; 1.
DR   Pfam; PF00289; CPSase_L_chain; 1.
DR   Pfam; PF02786; CPSase_L_D2; 1.
DR   PROSITE; PS00188; BIOTIN; 1.
DR   PROSITE; PS00867; CPSASE2; 1.
KW   Mitochondrion; Ligase; Biotin; ATP-binding; Transit peptide;
KW   Disease mutation; Polymorphism.
FT   TRANSIT       1     47       MITOCHONDRION (POTENTIAL).
FT   CHAIN        48    725       METHYLCROTONYL-COA CARBOXYLASE ALPHA
FT                                CHAIN.
FT   NP_BIND     209    214       ATP (POTENTIAL).
FT   ACT_SITE    339    339       BY SIMILARITY.
FT   BINDING     681    681       BIOTIN (BY SIMILARITY).
FT   DOMAIN      538    541       POLY-SER.
FT   DOMAIN      713    718       POLY-GLU.
FT   VARIANT     289    289       A -> V (in MCGI; mild form).
FT                                /FTId=VAR_012785.
FT   VARIANT     325    325       M -> R (in MCGI).
FT                                /FTId=VAR_012786.
FT   VARIANT     385    385       R -> S (in MCGI; severe form).
FT                                /FTId=VAR_012787.
FT   VARIANT     437    437       L -> P (in MCGI; severe form).
FT                                /FTId=VAR_012788.
FT   VARIANT     464    464       P -> H.
FT                                /FTId=VAR_012789.
FT   VARIANT     532    532       D -> H (in MCGI; severe form).
FT                                /FTId=VAR_012790.
FT   VARIANT     535    535       S -> F (in MCGI; asymptomatic form).
FT                                /FTId=VAR_012791.
FT   CONFLICT    469    469       F -> L (IN REF. 3 AND 4).
SQ   SEQUENCE   725 AA;  80433 MW;  B847C6B8O6O5B6C0 CRC64;
```

Query Match   15.7%; Score 911; DB 1; Length 725;
Best Local Similarity   36.5%;   Pred. No. 1.2e-41;
Matches 210; Conservative   96; Mismatches 204; Indels   66; Gaps   9;

```
Qy    4 HTSSTLPAFKKILVANRGEIAVRAFRAALFTGAATVAIYPREDRGSFHRSFASEAVRIGT 63
          :||: ||||||:||||||||  |:: ||| | :  : ||:  ||::|||| |:  ||
Db   41 YTTATGRNITKVLIANRGEIACRVMRTAKKLGVQTVAVYSEADRNSMHVDMADERAYSIGP 100

Qy   64 EGSPVKAYLDIDEIIGAAKKVKADAIYPGVGFLSENAQLARECAENGITFIGPTPEVLDL 123
          || :|| ::| || ||||| |||||| || ||:|| | |::||||||| |
Db  101 APSQ-QSYLSMEKIIQVAKTSAAQAIHPGCGFLSENMEFAELCKQEGIIFIGPPPSAIRD 159

Qy  124 TGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKS-AEGQTYPIFVKAVAGGGGRGMRFVA 182
          |||:| :|:::: ||::|| ::: |||:||:|||:|:|
Db  160 MGIKSTSKSIMAAAGVPVVEGYHGEDQSDQCLKEHARRIGYPVMIKAVRGGGGKGMRIVR 219

Qy  183 SPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVHLYERDCSLQ 242
          || |: :||| | :| | ||||| |: ::|| ||::|:|||| |:|| |||||:
Db  220 SEQEFQEQLESARREAKKSFNDDAMLIEKFVDTPRHVEVQVFGDHHGNAVYLFERDCSVQ 279

Qy  243 RRHQKVVETAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPR 302
          |||||:|||: |:  :  : :: |: |:::|| |||||||:| :|||::| :||:||||||
Db  280 RRHQKIIEEAPAPGIKSEVRKKLGEAAVRAAKAVNYVGAGTVEFIMDSKHNFCFMEMNTR 339

Qy  303 IQVEHTVTEEVTRVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFR 362
          :||: |||||:| : ||:|: | || |:|| | : : | :::|
Db  340 LQVEHPVTEMITGTDLVEWQLRIAAG--EKIPLSQEEITLQGHAFEARIYAEDPSNNFM 396
```

```
                GenCore version 5.1.6
        Copyright (C) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    September 24, 2003, 15:27:07 ; Search time 17.8668 Seconds
                                          (without alignments)
                                          2699.668 Million cell updates/sec

Title:          US-09-974-973a-19
Perfect score:  5788
Sequence:       1 MSTHTSSTLPAFKKILVANR.........RVVVPAATKVEGDLIVVVS 1140

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters:   328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_AA:*
             1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
             2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
             3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
             4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
             5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
             6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5788	100.0	1140	3	US-09-220-081-2	Sequence 2, Appli
2	5788	100.0	1140	3	US-09-677-575-2	Sequence 2, Appli
3	3483	60.2	1124	4	US-08-311-731A-10	Sequence 10, Appli
4	2464	42.6	1154	4	US-09-134-001C-3428	Sequence 3428, Ap
5	2174	37.6	973	4	US-09-107-532A-4810	Sequence 4810, Ap
6	1192	20.6	694	4	US-09-433-043B-126	Sequence 126, App
7	1083	18.7	456	4	US-09-634-238-276	Sequence 276, App
8	1043.5	18.0	593	4	US-09-433-043B-122	Sequence 122, App
9	1036.5	17.9	447	2	US-08-611-107-6	Sequence 6, Appli
10	1036.5	17.9	447	3	US-08-422-560A-6	Sequence 6, Appli
11	1036.5	17.9	447	3	US-08-468-793-6	Sequence 6, Appli
12	1004.5	17.4	453	2	US-08-611-107-8	Sequence 8, Appli
13	1004.5	17.4	453	2	US-08-422-560A-8	Sequence 8, Appli
14	1004.5	17.4	453	3	US-08-468-793-8	Sequence 8, Appli
15	1002.5	17.3	453	4	US-09-433-043B-121	Sequence 121, App
16	995.5	17.2	453	1	US-07-956-700B-6	Sequence 6, Appli
17	995.5	17.2	453	1	US-08-476-537-6	Sequence 6, Appli
18	995.5	17.2	453	1	US-08-485-607-6	Sequence 6, Appli
19	995.5	17.2	453	2	US-08-475-879-6	Sequence 6, Appli
20	995.5	17.2	453	4	US-09-433-043B-6	Sequence 6, Appli
21	967.5	16.7	474	4	US-09-328-352-7562	Sequence 7562, Ap
22	966.5	16.7	1116	4	US-09-252-991A-24374	Sequence 24374, A
23	952	16.4	605	4	US-09-433-043B-123	Sequence 123, App
24	938	16.2	448	5	US-08-074-121-3	Sequence 3, Appli
25	938	16.2	448	5	PCT-US94-06447-3	Sequence 3, Appli
26	937	16.2	427	1	US-07-956-700B-3	Sequence 3, Appli
27	937	16.2	427	1	US-08-476-537-3	Sequence 3, Appli
28	937	16.2	427	1	US-08-485-607-3	Sequence 3, Appli
29	937	16.2	427	2	US-08-475-879-3	Sequence 3, Appli
30	937	16.2	427	4	US-09-433-043B-3	Sequence 3, Appli
31	935.5	16.2	454	4	US-09-198-452A-197	Sequence 197, App
32	935.5	16.2	701	4	US-09-252-991A-27999	Sequence 27999, A
33	926	16.0	536	3	US-08-662-344-2	Sequence 2, Appli
34	924.5	16.0	465	1	US-09-252-991A-26980	Sequence 26980, A
35	924	16.0	449	1	US-08-074-121-6	Sequence 6, Appli
36	924	16.0	449	5	PCT-US94-06447-6	Sequence 6, Appli
37	901	15.6	676	4	US-09-252-991A-26143	Sequence 26143, A
38	899	15.5	670	4	US-09-328-352-6725	Sequence 6725, A
39	891	15.4	652	4	US-09-328-352-5587	Sequence 5587, Ap
40	889	15.4	573	4	US-09-328-352-6420	Sequence 6420, Ap
41	858.5	14.8	1212	4	US-09-252-991A-26616	Sequence 26616, A
42	851.5	14.7	461	4	US-09-134-001C-3604	Sequence 3604, Ap
43	843.5	14.6	722	4	US-09-433-043B-125	Sequence 125, APP
44	841	14.5	612	4	US-09-252-991A-19134	Sequence 19134, A
45	835.5	14.4	453	4	US-09-252-991A-19829	Sequence 19929, A

ALIGNMENTS

```
RESULT 1
US-09-220-081-2
; Sequence 2, Application US/09220081
; Patent No. 6171833
; GENERAL INFORMATION:
; APPLICANT: Sinskey, Anthony J.
; APPLICANT: Lessard, Philip A.
; APPLICANT: Willis, Laura B.
; APPLICANT: Stephanopoulos, Gregory
; TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
; FILE REFERENCE: 1533.0790000
; CURRENT APPLICATION NUMBER: US/09/220,081
; CURRENT FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 2
; LENGTH: 1140
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-220-081-2

Query Match             100.0%;  Score 5788;  DB 3;  Length 1140;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1140;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60

Qy   61 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120

Qy  121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180

Qy  181 VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240

Qy  241 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 300

Qy  301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360

Qy  361 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
```

```
Db  361  FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  421  AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV  480

Qy  481  TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  540

Qy  541  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  600

Qy  601  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  660

Qy  661  IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR  720

Qy  721  PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  780

Qy  781  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL  840

Qy  841  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHIVGAGVDPADF  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHIVGAGVDPADF  900

Qy  901  AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDA   960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDA   960

Qy  961  DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP  1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP  1020

Qy 1021  LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA  1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA  1080

Qy 1081  GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS  1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081  GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS  1140
```

```
US-09-677-575-2

Query Match        100.0%;   Score 5788;   DB 4;   Length 1140;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1140;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    1  MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR   60

Qy   61  IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV  120

Qy  121  LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF  180

Qy  181  VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS  240

Qy  241  LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN   300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN   300

Qy  301  PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG  360

Qy  361  FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL  420

Qy  421  AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV  480

Qy  481  TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT  540

Qy  541  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE  600

Qy  601  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA  660

Qy  661  IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR  720

Qy  721  PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS  780

Qy  781  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL  840

Qy  841  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHIVGAGVDPADF  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHIVGAGVDPADF  900

Qy  901  AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDA   960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDA   960

Qy  961  DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP  1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP  1020
```

```
Db   182 EAASREAESAFGDPTVYLEQAVINPRHIEVQILADNLGDVIHLYERDCSVQRRHQKVIEL 241
             :::|| ||||||| ||   ||||||||||:| |:|  |:|||||||| :||||||||||
Qy   252 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE 311
         || |||| ||| ::| ||||:|     ||  ||||||:|| |||||||||||| |||||||
Db   242 APAPHLDAELRYKMCVDAVAFARHIGYSCAGTVEFLLDERGEYVFIEMNPRVQVEHTVTE 301

Qy   312 EVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAY 371
         |:| |||||:|:|:|||:||::|||:|:|| ||||||||||||||||:||||  |||||
Db   302 EIITDVDLVASQLRIAAGETLEQLGLRQEDIAPHGAALQCRITTEDPANGFRP-TRAGSAR 360

Qy   372 RSPGGAGVRLDGAAQLGGEITAHFDSMLVKM----TCRGSDFETAVARAQRALAEFTVS 426
           |  ||           ::::        :::|| |    ||:||:||||:: ||||:|
Db   361 CDPPAVPVSAWTAAP----TWRRNQPVLRLHAGQADLSGRDLPTAVSRARRAIAEFRIR 415

Qy   427 GVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVNKPH 486
         ||:|||:||: :| |||| | |::|||:|| | ||  || :|||  | :||||||||||:
Db   416 GVSTNIPFLQAVLDDPFRAGRVTTSFIDERPQLLTARASADRGTKILNFLADVTVNNPY 475

Qy   487 GVRPKDVAAPIDKLPNIKDLPL---PRGSRDRLKLQGPAAFARDLREQDALAVTDTTFR 542
         |  |   ::|  ||| | ||   |       || ||| ||| | | |   |:|||||||
Db   476 GSRPSTI-YPDDKLP--DLDLRAAPPAGSKQRLVKLGPEGFARWLRESAAVGVTDTTFR 531

Qy   543 DAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGATYDVAMRFLFEDPWDRLDELR 602
         ||||||||||||:  || |   : |:||:|||||:|| |||||:|||:|||||:||:||
Db   532 DAHQSLLATRVRTSGLSRVAPYLARTMPQLLSVECWGGATYDVALRFLKEDPWERLATLR 591

Qy   603 EAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAID 662
          |||| ::|||||||||||||||||:| :|||:|||: ||||||||||||:: |:|||||
Db   592 AAMPNICLQMLLRGRNTVGYTPYPEIVISAFVQEATATGIDIFRIFDALNNIESMRPAID 651

Qy   663 AVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPA 722
         ||  | :|:|||||:|:|||:||  :|||||||||:||::| :|||:|||||||||||| 
Db   652 AVRETGSAIAEVAMCYTGDLTDPGEQLYTLDYYLKLAEQTVDAGAHVLAIKDMAGLLRPP 711

Qy   723 AVIKLVTALRLREEPDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLS 782
         |    ||:||  | ||||||:||||||.||||:| ||::.||||||||:||| ||||||:||
Db   712 AAQRLVSALRSRFDLPVHLHTHDTPGGQLASYVAAWHAGADAVDGAAAPLAGTTSQPALS 771

Qy   783 AIYAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSN 842
         :|:|| ||||  ||||||:|  |||||||:|:: ||||||:|||||||| ||||||||||
Db   772 SIVAAAAHTEYDTGLSLSAVCALEPYWEALRKVYAPFESGLPGPTGRVYHHEIPGGQLSN 831

Qy   843 LRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAA 902
         || ||:|||| |||| || :|:|.: :||:||||||:|||||||||:|||||| ::|||
Db   832 LRQQAIALGLGDRFEEIEEAYAGADRVLGRLVKVTPTSKVVGDLALALVGAGVSADEFAS 891

Qy   903 DPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDADD 962
         ||  : ||:|| :|||||||:|||||||||||:||.||: ::   |  |::::  | :
Db   892 DPARFGIPESVLGFLRGELGDPPGGWPEPLRTAALAGRGAAR-PTAQLAADDEIALSSVG 950

Qy   963 SKERRNSLNRLLFPKPTEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLL 1022
          |:|   |||||||.|| :| :|:| | |: |  : |||||  |:| ::|   ||  ||
Db   951 AK-RQATLNRLLFPSPTKEFNEHREAYGDTSQLSANQFFYGLRQGEEHVKL-ERGVELL 1008

Qy  1023 VRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAGV 1082
           |:|||||| :|||.|   ::||| ||:||||| : ||:  |||||:|  ||||||||
Db  1009 IGLEAISEPDERGMRTVWCILNGQLRPVLVRDRSIASAVPAAEKADRGNPHIAAPFAGV 1068

Qy  1083 VTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVV 1124
         |||.| :|  | |||.:|.|||||||| ||| | |:::||||
Db  1069 VTVGVCVGERVGAGQTIATIEAMKMEAPITAPVAGTVERVAV 1110
```

RESULT 4
US-09-134-001C-3428
; Sequence 3428, Application US/09134001C
; Patent No. 6680370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C

```
Qy  1021 LLVRLDAISEPPDKGMRNVVANVNGQIRPMRVRDRSVESVIATAEKADSSNKGHVAAPFA 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 LLVRLDAISEPPDKGMRNVVANVNGQIRPMRVRDRSVESVIATAEKADSSNKGHVAAPFA 1080

Qy  1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATTASVDGKIDRVVVPAATKVEGGDLIVVVS 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATTASVDGKIDRVVVPAATKVEGGDLIVVVS 1140
```

RESULT 3
US-08-311-731A-10
; Sequence 10, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1124 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM TUBERCULOSIS
US-08-311-731A-10

Query Match   60.2%; Score 3483; DB 4; Length 1124;
Best Local Similarity   62.2%; Pred. No. 7.7e-267;
Matches 698; Conservative 151; Mismatches 251; Indels 22; Gaps 9;

Qy   12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 71
         | |:|||||||||:|||||| ||| :||:|||:|||:| || | ||:||||| | ||||
Db   2 FSKVLVANRGEIAIRAFRAAYELGVGTVAVYPYEDRNSQHRLKADESYQIGDIGHPVHAY 61

Qy   72 LDIDEIIGAAKKVKADAIYPSYGFLSENAQLARECAENGITFIGPTPEVLLTGDKSRAV 131
         | :|||: ||::| |||||| ||||||| || ||| ||:|||: ||:| || |||:
Db   62 LSVDEIVATARRAGADAIYPGYGFLSENPDLAAACAAAGISFVGPSAEVLELAGNKSRAI 121

Qy   132 TAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKLA 191
          || :|||||| : :|| ::: :: | || |||:||||||||||||||| | ||:|: |
Db   122 AAAREAGLPVLMSSAPSASVDELLSVAAGMPFPLFVKAVAGGGGRGMRRVGDIAALPEAI 181

Qy   192 TEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVEI 251
         || |||| :||| ||:||:||||| :||||||.| ||:|:||||||||:||||||||:|:
```

```
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3428
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3428

Query Match    42.6%;   Score 2464;  DB 4;  Length 1154;
Best Local Similarity  45.4%;  Pred. No. 4e-186;
Matches  524;  Conservative 197; Mismatches 400;  Indels  34;  Gaps  14;

Qy     9 LPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPV 68
           |||:||| |||||||||:| :| |||| | || || :||| | ||:::| | :|::|
Db     6 LKQIKKLLVANRGEIAIRIFRAAAELNISTVAIYSNEDKSSLHRYKADESYLVGSDLGPA 65

Qy    69 KAYLDIDELIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKS 128
         ::||:: ||: :|| : ||||||||||||||:| || ||:||| |||  || |||| :
Db    66 ESYLNIERIIEVALRAGVDAIHPGYGFLSENEQFARRCAEEGIKFIGPHLEHLDMFGDKV 125

Qy   129 RAVTAAKKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRFVASPDEL 187
          : || | ||| :|  :  | :| |||||||  :||:|:||| || ||||| | ||| |
Db   126 KARTTAINANLPVIPGTDGPIESFEAAEQFANEAGYPLMIKATSGGGGKGMRIVRESSEL 185

Qy   188 RKLATEASREAFAAFGDGAVYVERAVINPQHIEVQILGDHTGEVHLYERDCSLQRRHQK 247
          : |: |:|||: ||| | || |||: ||:|||||:||| |  ||||||||:|||||||
Db   186 EDAFHRAKSEAEKSFGNSEVYIERYIDNPKHIEVQVIGDEFGNIIHLYERDCSVQRRHQK 245

Qy   248 VVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLV--DEKGNHVFIEMNPRIQV 305
         ||:::|| : :|:|||||| || | ::||| ||||||||| ||   :  |||||||:||
Db   246 VVFVAPSVGLSNKLRERICDAAIQLMENIKYVNAGTVEFLVSGDE---FFFIEVNPRVQV 302

Qy   306 EHTVTEEVTEVDLVKAQMRLAAGATL--KEIGLTQDKIKTHGAALQCRITTEDPNNGFR 362
         |||:||: |  |||| |:| |:|: |   :|   |:: :  | | ||||||||| |:|
Db   303 EHTITEMITGIDIVKTQILVANGESLFGDKISMPQQNEIQTLGYAIQCRITTEDPTNDFM 362

Qy   363 PDTGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALA 421
         ||:||| || |:|| ||||| | | | |||:  ||:||||: |  ::| |||| ||: |
Db   363 PDSGTIIAYRSSGGFGVRLDAGDGFQGAEISPYYDSLLVKLSTHAVSFKQAEEKMERSLR 422

Qy   422 EFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRIIDYLADVT 481
         |  : ||:|||:||: :|| | | :| ||  :|: | ||::|   : |:|  | |:|:|
Db   423 EMRIRGVKTNIPFLINVMRNDKFRSGDYTTKFIEETPELFDIAPTLDRGTKTLEYIGNVT 482

Qy   482 VNKPHGV--RPKDVAAPIDKLPNIKDLPLPR--GSRDRLKQLGPAAFARDLREQDALAVT 537
          | | |:   ||: | |:||:|| : : |    |:: : |::|  ::| : ||:||:| |
Db   483 INGFPNVEKRPKP-EYESTKIPKISQKKINQLFGTKQLIEQHGPTGVTNWVREQEDVLIT 541

Qy   538 DTTFRDAHQSLLATRVRSFALKPAAEBAVAKLTPELLSVEAWGGATYDVAMRPLFEDPWDR 597
         |||||||||||||||||:|| :::  ||:  ::|: |:|:||||||:|||  :|| ::|:|
Db   542 DTTFRDAHQSLLATRVRTKDMMNIASKTAEVFKDSFSLEMWGGATFDVAYNFLKENPWER 601

Qy   598 LDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQM 657
         |: || ||||  | |||| |  |||:|||:|| ::||:| |:|||:|||||| || | |
Db   602 LERLRKAIPNVLFQMLLRASNAVGYKNYPDNVIKKFVHESAKAGVDVFRIFDSLNWVDQM 661

Qy   658 RPAIDAVLETNTAVAEVAMAYSGD-LSDPNEKLYTDYYLKMAEEIVKSGAHILAIKDMA 716
          : :||: |:  | :|   :|:||  |: | :| : ||: ||:|:|:|||||||||||
Db   662 KVANEAVQEAG-MVSEGTICYTGDILNAERSNIYILDYYVMAKELEREGFHILAIKDMA 720

Qy   717 GLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTT 776
         |||:| | ::| :|| ::| ||:|:||||||:| |  |  :| :|||::| |:|: |||
Db   721 GLLKPKAAYELIGELREATHLPIHLHTHDTSGNGLLITYKQAIDAGVDIIDTPAVASMSGLT 780

Qy   777 SQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTRGVYRHEIP 836
         |||| :::| |  |  |: :| ::  ::| || ::|| :| :: : || |  |  ||:|
Db   781 SQPSANSLYYALNGFPRNLRTDIDGLEELSHYWSVVRPYYADFESDIKSPNTEIYQHEMP 840

Qy   837 GGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVD 896
         |||:||| :| :||| :||| :::|  ::|||:| |||||||||||||:|| ::| |||
Db   841 GGQYSNLSQQAKSLGLGERFDEVKEMYRRVNFLFGDLVKVTPSSKVVGDMALYMVQNDLD 900

Qy   897 PADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEQA 956
          ||  :|  | |:| ::: |: | |:  ||  :|:  :|  | :  :| ||| |   |
Db   901 EDTVINDGYKLDFPESVSSFFKGDIGQPVNGFNKKLQDVILKGQQ---PITERPGEYLE 956

Qy   957 HLDAD---------DSKERRNSLNRLLFPKPTEEFLEHRRRGNTSALDDREFFYGLV 1005
          :| :         :| : :::|:| :|    ::  :      |::||| :|||| |
Db   957 PVDFEAIRQELSDIQQDEVTEQDIISVLYPKVYKQYIQTKEQFGNVSLLDTPTFLFGMR 1016

Qy  1006 EGRETLIRLPDVRTPLIVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAE 1065
          | | :| | |  | |:::|:||||| |:|  |:  |:||| |  : :: :  ::|: |
Db  1017 NGETVEIEI-DTGKRLIIKLETISEPDENGKRTIYYAMNGQARRIYIQDENVKTNANVKP 1075

Qy  1066 KADSSNKGHVAAPFAGVVT-VTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVV 1124
         |||:||: |: |  | ||  | | ||||: | : : |||||||| ||:|  |||| : :|
Db  1076 KADKSNPNHIGAQMPGSVTEVKVSVGDEVQANQPLLITEAMKMETTIQAPFDGIIKQINV 1135

Qy  1125 PAATKVEGGDLIVVV 1139
             :||||||| | |
Db  1136 ANGDAIATGDLLVEI 1150
```

```
RESULT 5
US-09-107-532A-4810
; Sequence 4810, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
;  APPLICANT: Lynn A Doucette-Stamm  and David Bush
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;   ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;  NUMBER OF SEQUENCES: 7310
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;   STREET: 100 Beaver Street
;   CITY: Waltham
;   STATE: Massachusetts
;   COUNTRY: USA
;   ZIP: 02354
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE: CD/ROM ISO9660
;   COMPUTER: PC
;   OPERATING SYSTEM: <Unknown>
;   SOFTWARE: ASCII
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/107,532A
;   FILING DATE: 30-Jun-1998
;  PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 60/085,598
;   FILING DATE: 14 May 1998
;   APPLICATION NUMBER: 60/051571
;   FILING DATE: July 2, 1997
;  ATTORNEY/AGENT INFORMATION:
;   NAME: Ariniello, Pamela Deneke
;   REGISTRATION NUMBER: 40,489
;   REFERENCE/DOCKET NUMBER: GTC-012
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (781)893-5007
;   TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4810:
;  SEQUENCE CHARACTERISTICS:
;   LENGTH: 973 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;  MOLECULE TYPE: protein
;  HYPOTHETICAL: YES
;  ORIGINAL SOURCE:
;   ORGANISM: Enterococcus faecium
;  FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (B) LOCATION 1...973
;  SEQUENCE DESCRIPTION: SEQ ID NO: 4810:
```

; Sequence 126, Application US/09433043B
; Patent No. 6399342
; GENERAL INFORMATION:
;   APPLICANT: HASELKORN, ROBERT
;   APPLICANT: GORNICKI, PIOTR
;   TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-CoA CARBOXYLASE
;   FILE REFERENCE: ARCD:338US
;   CURRENT APPLICATION NUMBER: US/09/433,043B
;   CURRENT FILING DATE: 1999-10-25
;   PRIOR APPLICATION NUMBER: 08/475,879
;   PRIOR FILING DATE: 1995-06-07
;   PRIOR APPLICATION NUMBER: 07/956,700
;   PRIOR FILING DATE: 1992-10-02
;   NUMBER OF SEQ ID NOS: 128
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 126
;   LENGTH: 694
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Peptide
US-09-433-043B-126

US-09-107-532A-4810

Query Match   37.6%;   Score 2174; DB 4; Length 973;
Best Local Similarity   48.0%;   Pred. No. 2.8e-163;
Matches 454; Conservative 144; Mismatches 320; Indels   28; Gaps   8;

Qy   11 AFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKA 70
               : |||||||||||| ||| || ||: || | |:|| : |:
Db   11 SMKKVLVANRGEIAVRVFRACTELGIKTVGIYAEKNEYSVHRFKADEAYLVGQGKKPIDA 70

Qy   71 YLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRA 130
               ||:|: ||| :|: ||| ||||| |: || :|| | ||:|| ||||: |:
Db   71 YLDIEGIISIAKECGADAIHPGYGLLSENLNFAQRCEEKGIIFVGPKLHHLDIFGDKIKA 130

Qy   131 VTAAKKAGL-PVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRK 189
               : :|| :   :: || | || ::: :|| ||: | |:
Db   131 KAAAIEAGIASIPGTDGPIASIDDALEFAKQYGYPIMIKAALCGGGRGMRVAHDEKSARE 190

Qy   190 LATEASREAEAAFGDGAVYYERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVV 249
               : | | |:|| :||||::|: :|| | :| |||||:|||||:|||||
Db   191 GYERAKSEAKAAFGSDEVVVEKYIANPKHIEVQILGDTHGNVIHLFERDCSVQRRHQKVV 250

Qy   250 EIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTV 309
               |:|| :|| |||||||||||||||||||||||||||||||:||||||||||||||||||
Db   251 EVAPCVSMNEQQRQKICQAAVPLMKVGYVNARTVEFLV-EGDDLYFIEVNPRVQVEHTI 309

Qy   310 TEEVTEVDLVKAQMRLAAGATL-KELGLTQDK-IKTHGAALQCRITTEDPNNGFRPDTGT 367
               ||::|| || || :| ||: | |:| | :||||:||||||:||| :|||:|||
Db   310 TEMITDIDIVTTQLLIAQGLDLHKEIGLPQQEGIKLNGSAIQCRITTEDPLNNFLPDTGK 369

Qy   368 ITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVS 426
               || |||||| ||||| : | ||:|||||||| | :| || ||:| |: | |:
Db   370 IDTYRSPGGFGVRLDVGNAYAGYVTPYFDSLLVKVCTHGATFETAIQKMERCLKEFRIR 429

Qy   427 GVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVNKPH 486
               ||: |||||| :|:|: | :| | | | || : || :| || | ||:|
Db   430 GVKTNIPFMLNVITHPEFQSGNAKTFIDSTATLFEFPRLRDRGNKTMKYIGETTVNGFP 489

Qy   487 GVR--------PKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDAL 534
               |: ||||||||||||||||||:||||||||||||||||||||||:||
Db   490 GIESGEKPFYEEPRMPKDLITRTDYVT---------AKNVLDAKGADALVEWIKGQENL 539

Qy   535 AVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDP 594
               ||| ||||||||||||||||:| | ||||||||:|| ||||| |||| | || ||
Db   540 LLTDTTFRDAHQSLLATRVRTKDFKQIARLTGEGLPELFSSEMWGGATFDVAYRFLNEDP 599

Qy   595 WDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDV 654
               |||::|| | || || |||:||::|| | | | | :||||||||||||||||||||
Db   600 WQRLKIRSLMPNTLLQMLFRGSNAVGYSNYPDNVLVEFVKEAAAQGIDVFRIFDSLNWI 659

Qy   655 SQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKD 714
               |||| ||| | | ||| | |||||||||||||||| ||||||||||||||||||||||
Db   660 PQMEKSIQAVRDTG-KIAEAAICVTGDINDPSRAKYNVQYYKDMAKELEQLGAHIIAIKD 718

Qy   715 MAGLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSG 774
               |||||||||||||| ||:| | :||||| :||:| | | :||| |||| ||:|
Db   719 MAGLLKPQAAYRLISELKETTDLPIHLHTHDTSGNGIITYSAASKAGVDIVDVAMSAMSG 778

Qy   775 TTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHE 834
               |||||||||| |:|||:: :| | || |:| : ||: |: || ::
Db   779 NTSQPSMSSLYYALVNGPRLPEITIENAQKLNHYWEDVRMYYKPFENGLNAPETEVYMHE 838

Qy   835 IPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVIPSSKVVGDLALHLVGAG 894
               |||| :|| || ||:: |::| |::| | :| | |::|||||||| |:|:||||
Db   839 MPGGQYSNLQQQAKAVGLGHRWDEIKQMYHTVNLMFGDIVKVTPSSKVVGDMALFMVQND 898

Qy   895 VDPADFAADPQKIYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGR 940
               :| | :| || | | || | ||||| || | | | | | ||||
Db   899 LTEEDIYEKGETLSFPESVTFFQGELGQPVGGFPKKLQKIILKGR 944

RESULT 6
US-09-433B-126

Query Match   20.6%;   Score 1192; DB 4; Length 694;
Best Local Similarity   28.0%;   Pred. No. 9.8e-86;
Matches 317; Conservative 100; Mismatches 240; Indels 476; Gaps   11;

Qy   14 KILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEG--SPVKAY 71
               |||||||||||| |||||||||||||||||||||||||||||||||||||| :|| ||
Db   21 KILVANRGEIP IRIFRTAHELSMQTVAIYSHEDRLSTHKQKADEAYVIGEVGQYTPVGAY 80

Qy   72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 131
               |:|:| || : || | ||||||||| :| ||| |: |||| ||: | |::
Db   81 LAIDEIISIAQKHQVDFIHPGYGFLSENSEFADKVVKAGITWIGPPAEVIDSVGDKVSAR 140

Qy   132 TAAKKAGLPVLAEST-PSKNIDEIVKSAEGQTYPIFVKAVAGGGGRMRFVASPDELRKL 190
               || |||:::   : || |:||:||||||| ||:||||| || :|||| |: | |
Db   141 NLAAKANVPTVPGTPGPIETVEEALDFVNEYGYPVIIKAAFGGGGRGMRVVRGGDDVADA 200

Qy   191 ATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE 250
               | ||| ||||||: |||:|:|| | ||||||||| |||||||||||||:||||||||
Db   201 FQRATSEARTAGFNGTCFVERFLDKPKHIEVQLLADNHGNVHLFERDCSVQRRHQKVVE 260

Qy   251 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 310
               :|| :|| ||||||||:|||:||| |: | ||| |:|:| |:|||||||||||:|:
Db   261 VAPAKTLPREVRDAILTDAVKLAKECGYRNAGTAEFLVDNQNRHYFIEINPRIQVEHTIT 320

Qy   311 EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 370
               |||:||||||| :| || |:|||:|| |:|||:||||||||:||| :|||||||| ||
Db   321 EEITGIDIVAAQIQ-AAGASLPQLGLFQDKITTRGFAIQCRITTEDPAKNFQPDTGRIEV 379

Qy   371 YRSPGGAGVRLDGA-AQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVA 429
               |||| || ||| : ||: :|||||||:|||:||:|: |||| || ||:| || ||| |
Db   380 YRSAGGNGVRLDGNAYAGTIISPHYDSMLVKCSCSGSTYEIVRRKMIRALLEFRIRGVK 439

Qy   430 TNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVNKPHGVR 489
               ||| :| || |: | |
Db   440 TNIPFLLTLLTNPVFIEGTYWGTFIDDTPQLFQMVSSQNRAQKLLHYLADVA-------- 491

Qy   490 PKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDAHQSLL 549
Db   492 -- 491

Qy   550 ATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRLFEDPWDRLDELREAMPNVN 609
Db   492 -- 491

Qy   610 IQMLLRGRNIVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETNT 669
Db   492 -- 491

Qy   670 AVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLVT 729
Db   492 -- 491

```
Db   131 AGADLFADLHLPQQDALRENGAAIQCRITTEDPENNFMPDTGTINTYRSPGGFGIRLDVG  190
             |||  |  ||| :: | ||  ||||| | ::  |: |||| ||| :|||  :|||| |:
Qy   384 AAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREED  443
             :   :  ||||||:|   ::::: :|||||  |||| ||  :| | : | |||||||||||
Db   191 NAYAGAVVSPYFDSLLVKASVHAPSFPAAVAKMQRALHEFQITGVKTNVAFLEHLLATQT  250
            ||  : |::|| | :|| | |:||  ||  | || |:| :|  | |::| ||| :|| :|
Qy   444 FTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVNKPHGV-RPKDVAAPIDKLPN  502
           |  ||| | ||| | ||| |:  | ||| ||  | || || ::|  ||:: | | ||
Db   251 FRTGEAETAFIDAHPELLQVQAKPDIASRLLWYISDVTVNGFKGVERQSQKYYP--ELQY  308
             |  | | |: ||   |: ||| |||    ||| :|| |:  | |||::::|:  | |
Qy   503 IKDLPLPRGSRD---RLKQLGPAAFARDLREQDALAVTDTTFPRDAHQSLLATRVRSFALK  559
            :: | | ||||    |: :|  || ::|  |  | | || ||||||||| | || :| ||
Db   309 TRHFAAAKPQTDIVALLKNEGAQAVTDWVKAHPALLLTDTTFRDAHQSLFATRMRTRDML  368
            | | | |||  |:  :  | | || : |  |:| :|||| ||  || :| ||| :|::||
Qy   560 PAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNT  619
           | | | | | | ||| | ||||| ||| |||:|||| ||::| | |||:| ||| ||:|
Db   369 TVAEDMGNGLPNLFSMEVWGGATFDVAYRFLNEDPWVRLKKLRAALPHTLLQMLFRGSNA  428
              |  :: | ||| : |||||||||||| ||||:||| |: ||| | |::|||| ||:|
Qy   620 VGTPYPDSVCRAFVKEAASSGVDIFRI  647
           ||  : |:|  ||: :||::|||:||
Db   429 VGYQNYPDNVIKAFINQAANDGVDVFRI  456
```

```
Db   492 ----------------------------------------  491
Qy   730 ALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFA  789
Db   492 ----------------------------------------  491
Qy   790 HTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSNLRAQATA  849
Db   492 ----------------------------------------  491
Qy   850 LGLADRFELIEDNYAAVNEMLGRP-TKVTPSSKVVGDLALHLVGAGVDPADFEAADPQKYD  908
             | : :: :   :|| | |: :|    |:
Db   492 ----------------DNGSSIKGQIGLPKLKSNPSV-------------  517
Qy   909 IPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDADDSKERRN  968
Db   518 ----------------------------------------  517
Qy   969 SLNRLLFPKPIEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLRPDVRTPLLVRLDAI 1028
              :|:  | | ||| : || :: ||  |: :  |:  ||||| |||: |:|| ||:|||
Db   518 -----MYPRVYEDFQKMRETYGDLSVLPTRSFLSPLETDEELEVVIEQGKT-LIIKLQAV  571
Qy  1029 SEPDDK-GMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAGV-VTVT 1086
            : : |: ||||  |:| |:|   | ||||| :| :   |::| | | |||| |||| ||:|
Db   572 GDLNKKTGEREVYFDLNGEMRKIRVADRSQKVETVTKSKADMHDPLHIGAPMAGVIVEVK  631
Qy  1087 VAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVV        1139
            |::| |:|||:| || ||||||: || | ||:  :|  ||  |:||||:|:|
Db   632 VHKGSLIKKGQPVAVLSAMKMEMIISSPSDGQVKEVFVSDGENVDSSDLLVLL         684
```

```
Query Match             18.7%;  Score 1083;  DB 4;  Length 456;
Best Local Similarity   51.3%;  Pred. No. 2.1e-77;
Matches   66;  Conservative  66;  Mismatches 142;  Indels  10;  Gaps  7;
```

```
Qy   207 VYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVETAPAQHLDPELRDRIC  266
            :||| || ||:||||:||| |||:||||:|||||:|||:|||||||||| || | ||:||
Db    12 MYVEKYIASAKHVEVQVLGDEHGHLHLFERDCSVQRRQQKVVETAPAVALPLALRNRIC   71
Qy   267 ADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEVTEVDLVKAQMRLA  326
            | ||: |||: |:||||||||||| : |||: |||:|:|||| :|  ||::| |:  |:|
Db    72 QSAVDLMASLHYENAGTVEFLVD-GDQYYFIEVNPRVQVEHTIFELITGVDIVQSQLRIA  130
```

```
Query Match             18.0%;  Score 1043.5;  DB 4;  Length 593;
Best Local Similarity   44.6%;  Pred. No. 4.4e-74;
Matches   229;  Conservative  81;  Mismatches 180;  Indels  23;  Gaps  5;
```

```
Qy    12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY   71
          | :||:||||||| |  |  :  || || |  :  | ::||| ||: || | | |||  |
Db     3 FDKILIANRGEIALRILRACEEMGIATIAVHSTVDRNALHVQLADEAVCIG-EPASAKSY   61
Qy    72 LDIDEIIGAAKKVKADAIYPGVGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV  131
          |:| :|: || : ||| |:||:||||||| :| |||::|:||||||| | : ||||:: |
Db    62 LNIPNIIAAALTRNASAIHPGYGFLSENAKFAEICADHIAFIGPTPEAIRLMGDKSTAK  121
Qy   132 TAAKKAGLPVLAESTPSKNID-EIVKSAEGQTYPIFVKAVAGGGRGMRFVASPDELRKL  190
           | :||| | | | | ::  : |||:: || |||:|:||||||| ||||:| |||| :||
Db   122 ETMQKAGVPTVPGSEGLVETEQGLELAKDIGYPVMIKATAGGGGRGMRLVRSPDEFVKL  181
Qy   191 ATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE  250
           |  :  || ||| : | :||:: : |:|||:|||| ||:::| | |||||:|||:|| :|
Db   182 FLAAQGEAGAAFGNAGVYIEKFIERPRHIEFQILADNYGNVIHLGERDCSIQRRNQKLLE  241
Qy   251 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT  310
           :||: ||| ::| : :  | ||| | || |||||||:| :| ::|:|||: |||||| ||
Db   242 EAPSPALDSDLREKMGQAAVKAAQFINYAGAGTIEFLLDRSGQFYFMEMNTRIQVEHPVT  301
```

```
Qy  12  FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY  71
        | |||:||||||:|  :| ||: | ||::::   ||  || ::||  |  | ::| .:
Db   3  FDKILIANRGEIAIRLRILRACEEMGIATIAVHSTVDRNALHVQLADEAVCIG-EPASAKSY  61

Qy  72  LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV  131
        |:| ||| | ||:|||:||||||||||| :: | ||::| :||||||| : | |||| |
Db  62  LNIPNIIAALTRNASAIHPGYGFLSENAKFAEICADHHIAFIGPTPEAIRLMGDKSTAK  121

Qy 132  TAAKKAGLPVLAESTPSKNID-EIVKSAEGQTYPIFVKAVAGGGRGMRFVASPDELRKL  190
        : |:|||:| : :      :|| : : | | |||::|||| || ||||| |:|||| ||
Db 122  ETMQKAGVPTVPGSEGLVETEQGLELAKDIGYPVMIKATAGGGGRGMRLVRSPDEFVKL  181

Qy 191  ATEASREAAFGDGAVVVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE  250
        | ::|||||||| |||:|:| ||| || |||| :: | |:|| |||| |||::|| :
Db 182  FLAAQGEAGAAFGNAGVVIEKFIERPRHIEFQLLADNYGNVIHLGERDCSIQRRNQKLLE  241

Qy 251  IAPAQHLDPELRDRICADAVKFCRSIGYQGAGVTVEFLVDEKGNHVFIEMNPRIQVEHTVT  310
        ||  :::| | ::   ||||||| || || || |||| | :||::|::|||:||||| ||
Db 242  EAPSPALDSDLREKMGQAAVKAAQFINYTGAGTIEFLLDRSGQFYFMEMNTRIQVEHPVT  301

Qy 311  EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA  370
        | || ||| |:: |||:|| ||: ||||::: | || ::||| ||| | ||| || :||
Db 302  EMVTGVDLLVEQIRIAQGERLR--LTQDQVVLRGHAIECRINAEDPDHDFRPAPGRISG  358

Qy 371  YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT  430
        |  |||:|||:| :  :  |||||||:|:  : :|  |: : | | |||| | |:|| |
Db 359  YLPPGGPGVRIDSHVYTDYQIPPYYDSLIGKLIVWGPDRATAINRMKRALRECAITGLPT  418

Qy 431  NIGFLRALLREEDFTSKRIATGFIAD  456
        .||| :|: :::| : || |:|:|:|
Db 419  TIGFHQRIMENPQFLQGNVSTSFVQE  444
```

RESULT 10
US-08-422-560A-6
; Sequence 6, Application US/08422560A
; Patent No. 5910626
; GENERAL INFORMATION:
;   APPLICANT: Haselkorn, Robert
;   APPLICANT: Gornicki, Piotr
;   TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
;   TITLE OF INVENTION: METHODS FOR USE
;   NUMBER OF SEQUENCES: 31
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Arnold, White & Durkee
;   STREET: P.O. Box 4433
;   CITY: Houston
;   STATE: TX
;   COUNTRY: USA
;   ZIP: 77210-4433
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/422,560A
;   FILING DATE: 14-APR-1995
;   CLASSIFICATION: 800
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/956,700
;   FILING DATE: 02-OCT-1992
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Wilson, Mark B.
;   REGISTRATION NUMBER: 37,259
;   REFERENCE/DOCKET NUMBER: ARCD:152/WIM
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 512-418-3000
;   TELEFAX: 512-474-7577
; INFORMATION FOR SEQ ID NO: 6:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 447 amino acids
;   TYPE: amino acid

```
Qy 311  EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA  370
        | || ||| |:: |||:|| ||: ||||::: | || ::||| ||| | ||| || :||
Db 302  EMVTGVDLLVEQIRIAQGERLR--LTQDQVVLRGHAIECRINAEDPDHDFRPAPGRISG  358

Qy 371  YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT  430
        |  |||:|||:| :  :  |||||||:|:  : :|  |: : | | |||| | |:|| |
Db 359  YLPPGGPGVRIDSHVYTDYQIPPYYDSLIGKLIVWGPDRATAINRMKRALRECAITGLPT  418

Qy 431  NIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYIA------------  478
        .||| :|: :::| : || |:|:|:|  :| :| :  |:   | ::|
Db 419  TIGFHQRIMENPQFLQGNVSTSFVQE----MNKPLDFNEIRQLLITIAQTDIAEVTLKSD  474

Qy 479  --DVTVNKPHGVRPKDVAAPIDKLPNIKDLPLP  509
          :: :|: : |: |  ||:  |    |||
Db 475  DFELIVRKAVGVNNSVPVVTAPLSGVGVSGGLP  507
```

RESULT 9
US-08-611-107-6
; Sequence 6, Application US/08611107
; Patent No. 5801233
; GENERAL INFORMATION:
;   APPLICANT: Haselkorn, Robert
;   APPLICANT: Gornicki, Piotr
;   TITLE OF INVENTION: NUCLEIC ACID COMPOSITIONS ENCODING
;   TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE AND USES
;   TITLE OF INVENTION: THEREFOR
;   NUMBER OF SEQUENCES: 40
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Arnold, White & Durkee
;   STREET: P.O. Box 4433
;   CITY: Houston
;   STATE: Texas
;   COUNTRY: United States of America
;   ZIP: 77210
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/611,107
;   FILING DATE: Concurrently Herewith
;   CLASSIFICATION: 800
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US SN 07/956,700
;   FILING DATE: 02-OCT-1992
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US PCT/US93/09340
;   FILING DATE: 30-SEP-1993
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US SN 08/422,560
;   FILING DATE: 14-APR-1995
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Kitchell, Barbara S.
;   REGISTRATION NUMBER: 33,928
;   REFERENCE/DOCKET NUMBER: ARCD:221
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (512) 418-3000
;   TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 6:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 447 amino acids
;   TYPE: amino acid
;   STRANDEDNESS:
;   TOPOLOGY: linear
US-08-611-107-6

Query Match   17.9%; Score 1036.5; DB 1; Length 447;
Best Local Similarity 48.4%; Pred. No. 1e-73;
Matches 216; Conservative 73; Mismatches 152; Indels 5; Gaps 3;

```
;     STRANDEDNESS:
;     TOPOLOGY:  linear
US-08-422-560A-6

Query Match             17.9%;  Score 1036.5;  DB 2;  Length 447;
Best Local Similarity   48.4%;  Pred. No. 1e-73;
Matches  216;  Conservative  73;  Mismatches  152;  Indels  5;  Gaps  3;

Qy   12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 71
           | :||:|||||:||| || ||||:::  || ::| |||| | :||| ||| |:|
Db    3 FDKILIANRGEIALRILRACEEMGIATIAVHSTVDRNALHVQLADEAVCIG-EPASAKSY 61

Qy   72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 131
          |:| ||  :|| :|||||||||||||||||  ||| |||||:| |: |||| ||||
Db   62 LNIPNIIAAALTRNASAIHPGYGFLSENAKFAEICADHHIAFIGPTPEAIRLMGDKSTAK 121

Qy  132 TAAKKAGLPVLAESTPSKNID-EIVKSAEGQTYPIFVKAVAGGGRGMRFVASPDELRKL 190
          :||||||||| :||| ||||| ||||||| ||  :||| ||||||||||| :|| ||
Db  122 ETMQKAGVPVTVPGSEGLVETEQEGLELAKDIGYPVMIKATAGGGRGMRLVRSPDEFVKL 181

Qy  191 ATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE 250
           | ||| || |  || |||||:|| |||||:|:| |||||||| ||||:|::||:::|
Db  182 FLAAQGEAGAAFGNAGVYIEKFIERPRHIEFQILADNYGNVIHLGERDCSIQRRNQKLLE 241

Qy  251 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 310
          ||| :| | ||  | :| | |||  ||| ||||||||:||  |||||||| |||||| |
Db  242 EAPSPALDSDLREKMGQAAVKAAQFINYTGAGTIEFLLDRSGQFYFMEMNTRIQVEHPVT 301

Qy  311 EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGITA 370
          | :|| | :|:::| |  || | |:|  | ||||  |   || |: ||  | | :
Db  302 EMVTGVDLLVEQIRIAQGERLR--LTQDQVVLRGHAIECRINAEDPDHDFRPAPGRISG 358

Qy  371 YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 430
          | |||| |||||| |||  :|:| || ||:|  | |  |||||:| ::|: :||||||
Db  359 YLPPGGPGVRIDSHVYTDYQIPPYYDSLIGKLIVWGPDRATAINRMKRALRECAITGLPT 418

Qy  431 NIGFLRALLREEDFTSKRIATGFIAD 456
          ||| || ||||  :|||| |||
Db  419 TIGFHQRIMENPQFLQGNVSTSFVQE 444

RESULT 11
US-08-468-793-6
; Sequence 6, Application US/08468793
; Patent No. 6177267
; GENERAL INFORMATION:
;   APPLICANT:  Haselkorn, Robert
;   APPLICANT:  Gornicki, Piotr
;   TITLE OF INVENTION:  ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
;   TITLE OF INVENTION:  METHODS OF USE
;   NUMBER OF SEQUENCES:  29
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Arnold, White & Durkee
;     STREET:  P.O. Box 4433
;     CITY:  Houston
;     STATE:  Texas
;     COUNTRY:  United States of America
;     ZIP:  77210
; COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Floppy disk
;   COMPUTER:  IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS/ASCII
;   SOFTWARE:  PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/468,793
;   FILING DATE:  06-JUN-1995
;   CLASSIFICATION:  800
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  US 08/422,560
;   FILING DATE:  14-APR-1995
;   APPLICATION NUMBER:  US SN 07/956,700
;   FILING DATE:  02-OCT-1992
;   CLASSIFICATION:  800
```

```
;   APPLICATION NUMBER:  PCT/US93/09340
;   FILING DATE:  30-SEP-1993
;   CLASSIFICATION:  800
; ATTORNEY/AGENT INFORMATION:
;   NAME:  Kitchell, Barbara S.
;   REGISTRATION NUMBER:  33,928
;   REFERENCE/DOCKET NUMBER:  ARCD:152/KIT
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  (512) 418-3000
;   TELEFAX:  (713) 789-2679
;   TELEX:  79-0924
; INFORMATION FOR SEQ ID NO:  6:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  447 amino acids
;     TYPE:  amino acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  protein
US-08-468-793-6

Query Match             17.9%;  Score 1036.5;  DB 3;  Length 447;
Best Local Similarity   48.4%;  Pred. No. 1e-73;
Matches  216;  Conservative  73;  Mismatches  152;  Indels  5;  Gaps  3;

Qy   12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 71
           | :||:|||||:||| || ||||:::  || ::| |||| | :||| ||| |:|
Db    3 FDKILIANRGEIALRILRACEEMGIATIAVHSTVDRNALHVQLADEAVCIG-EPASAKSY 61

Qy   72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 131
          |:| ||  :|| :|||||||||||||||||  ||| |||||:| |: |||| ||||
Db   62 LNIPNIIAAALTRNASAIHPGYGFLSENAKFAEICADHHIAFIGPTPEAIRLMGDKSTAK 121

Qy  132 TAAKKAGLPVLAESTPSKNID-EIVKSAEGQTYPIFVKAVAGGGRGMRFVASPDELRKL 190
          :||||||||| :||| ||||| ||||||| ||  :||| ||||||||||| :|| ||
Db  122 ETMQKAGVPVTVPGSEGLVETEQEGLELAKDIGYPVMIKATAGGGRGMRLVRSPDEFVKL 181

Qy  191 ATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE 250
           | ||| || |  || |||||:|| |||||:|:| |||||||| ||||:|::||:::|
Db  182 FLAAQGEAGAAFGNAGVYIEKFIERPRHIEFQILADNYGNVIHLGERDCSIQRRNQKLLE 241

Qy  251 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 310
          ||| :| | ||  | :| | |||  ||| ||||||||:||  |||||||| |||||| |
Db  242 EAPSPALDSDLREKMGQAAVKAAQFINYTGAGTIEFLLDRSGQFYFMEMNTRIQVEHPVT 301

Qy  311 EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGITA 370
          | :|| | :|:::| |  || | |:|  | ||||  |   || |: ||  | | :
Db  302 EMVTGVDLLVEQIRIAQGERLR--LTQDQVVLRGHAIECRINAEDPDHDFRPAPGRISG 358

Qy  371 YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 430
          | |||| |||||| |||  :|:| || ||:|  | |  |||||:| ::|: :||||||
Db  359 YLPPGGPGVRIDSHVYTDYQIPPYYDSLIGKLIVWGPDRATAINRMKRALRECAITGLPT 418

Qy  431 NIGFLRALLREEDFTSKRIATGFIAD 456
          ||| || ||||  :|||| |||
Db  419 TIGFHQRIMENPQFLQGNVSTSFVQE 444

RESULT 12
US-08-611-107-8
; Sequence 8, Application US/08611107
; Patent No. 5801233
; GENERAL INFORMATION:
;   APPLICANT:  Haselkorn, Robert
;   APPLICANT:  Gornicki, Piotr
;   TITLE OF INVENTION:  NUCLEIC ACID COMPOSITIONS ENCODING
;   TITLE OF INVENTION:  ACETYL-CoA CARBOXYLASE AND USES
;   TITLE OF INVENTION:  THEREFOR
;   NUMBER OF SEQUENCES:  40
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Arnold, White & Durkee
;     STREET:  P.O. Box 4433
;     CITY:  Houston
;     STATE:  Texas
```

COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,107
FILING DATE: Concurrently Herewith
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US SN 07/956,700
FILING DATE: 02-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US93/09340
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US SN 08/422,560
FILING DATE: 14-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: ARCD:221
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 453 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-611-107-8

Query Match   17.4%;   Score 1004.5; DB 1; Length 453;
Best Local Similarity   47.1%;   Pred. No. 3.5e-71;
Matches 209; Conservative 72; Mismatches 158; Indels 5; Gaps 3;

Qy 12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 71
Db 3   FNKILIANRGEIALRILRTCEELGIGTIAVHSTVDRNALHVQLADEAVCIG-EAASSKSY 61

Qy 72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 131
Db 62 LNIPNIIAAALTRNASAIHPGYGFLAENARFAEICADHHLTFIGPSPDSIRAMGDKSTAK 121

Qy 132 TAAKKAGLPVLAESTP-SKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKL 190
Db 122 ETMQRVGVPTIPGSDGLLTDVDSAAKVAAEIGYPVMIKATAGGGGRGMRLVREPADLEKL 181

Qy 191 ATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE 250
Db 182 FLAAQGEAEAAFGNPGLYLEKFIDRPRHVEFQILADAYGNVVHLGERDCSIQRRHQKLLE 241

Qy 251 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 310
Db 242 EAPSPALSADLRQKMGDAAVKVAQAIGYIGAGTVEFLVDATGNFYFMEMNTRIQVEHPVT 301

Qy 311 EEVTEVDLVKAQMRLAAGATIKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 370
Db 302 EMITGLDLIAEQIRIAQGEALR--FRQADIQLRGHAIECRINAEDPEYNFRPNPGRITG 358

Qy 371 YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 430
Db 359 YLPPGPGVRVDSHVTDYEIPPYYDSLIGKLIVWGATREEAIARMQRALRECAITGLPT 418

Qy 431 NIGFLRALLREEDFTSKRIATGFI 454
Db 419 TLSFHQLMLQMPEFLRGELYTNFV 442

RESULT 13

US-08-422-560A-8
Sequence 8, Application US/0842256OA
Patent No. 5910626
GENERAL INFORMATION:
APPLICANT: Haselkorn, Robert
APPLICANT: Gornicki, Piotr
TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
TITLE OF INVENTION: METHODS FOR USE
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,560A
FILING DATE: 14-APR-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/956,700
FILING DATE: 02-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mark B.
REGISTRATION NUMBER: 37,259
REFERENCE/DOCKET NUMBER: ARCD:152/WIM
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-418-3000
TELEFAX: 512-474-7577
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 453 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-422-560A-8

Query Match   17.4%;   Score 1004.5; DB 2; Length 453;
Best Local Similarity   47.1%;   Pred. No. 3.5e-71;
Matches 209; Conservative 72; Mismatches 158; Indels 5; Gaps 3;

Qy 12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 71
Db 3   FNKILIANRGEIALRILRTCEELGIGTIAVHSTVDRNALHVQLADEAVCIG-EAASSKSY 61

Qy 72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 131
Db 62 LNIPNIIAAALTRNASAIHPGYGFLAENARFAEICADHHLTFIGPSPDSIRAMGDKSTAK 121

Qy 132 TAAKKAGLPVLAESTP-SKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKL 190
Db 122 ETMQRVGVPTIPGSDGLLTDVDSAAKVAAEIGYPVMIKATAGGGGRGMRLVREPADLEKL 181

Qy 191 ATEASREAEAAFGDGAVVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE 250
Db 182 FLAAQGEAEAAFGNPGLYLEKFIDRPRHVEFQILADAYGNVVHLGERDCSIQRRHQKLLE 241

Qy 251 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 310
Db 242 EAPSPALSADLRQKMGDAAVKVAQAIGYIGAGTVEFLVDATGNFYFMEMNTRIQVEHPVT 301

Qy 311 EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 370
Db 302 EMITGLDLIAEQIRIAQGEALR--FRQADIQLRGHAIECRINAEDPEYNFRPNPGRITG 358

Qy 371 YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 430

Db 122 ETMQRVGVPTIPGSDGLLTDVDSAAKVAAEIGYPVMIKATAGGGRGMRLVREPADLEKL 181

Qy 191 ATEASREAAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE 250
         ||:||||||| ||||||||:||||:||| |:|||||||| :|| ||||||||||||:|
Db 182 FLAAQGEAEAAFGNPGLYLEKFIDRPRHVEFQILADAYGNVHLGERDCSIQRRHQKLLE 241

Qy 251 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 310
         || :| ||||| ||| ||| | ||:||| ||||||||||| ||| |||||||||||| |
Db 242 EAPSPALSADLRQKMGDAAVKYAQAIGYIGAGTVEFLVDATGNFYFMEMNTRIQVEHPVT 301

Qy 311 EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 370
         | || | |:|:|:| ::|| :: | |
Db 302 EMITGLDLIAEQIRIAQGEALR--FRQADIQLRGHAIECRINAEDPEYNFRPNPGRITG 358

Qy 371 YRSPGGAGVRLDGAAQLGGETTAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 430
         |:|| ||||||| | ||||| || || :| |:||||| |||:| | |:|: :|||
Db 359 YLPPGGPGVRDSHVTDYEIPPYYDSLIGKLIVWGATREEAIARMQRALRECAITGLPT 418

Qy 431 NIGFLRALLREEDFTSKRIATGFI 454
         :|:|::| :||| || ||||||::
Db 419 TLSFHQLMLQMPEFLRGELYINFV 442

RESULT 15
US-09-433-043B-121
; Sequence 121, Application US/09433043B
; Patent No. 6399342
; GENERAL INFORMATION:
; APPLICANT: HASELKORN, ROBERT
; APPLICANT: GORNICKI, PIOTR
; TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-CoA CARBOXYLASE
; FILE REFERENCE: ARCD:338US
; CURRENT APPLICATION NUMBER: US/09/433,043B
; CURRENT FILING DATE: 1999-10-25
; PRIOR APPLICATION NUMBER: 08/475,879
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 07/956,700
; PRIOR FILING DATE: 1992-10-02
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 121
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-433-043B-121

Query Match   17.3%; Score 1002.5; DB 4; Length 453;
Best Local Similarity 47.1%; Pred. No. 5e-71;
Matches 209; Conservative 72; Mismatches 158; Indels 5; Gaps 3;

Qy   12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 71
         ||:||||||||||||||||:|| |:: | |:|: | |||: |||||||||:||||||||
Db   3 FNKILIANRGEIALRILRTCEELGIGTIAVHSTVDRNALHVQLADEAVCIG-EAASSKSY 61

Qy   72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 131
         |:||||||||:|||||||||||||:|||:|||| | :||:||||:||||:||||||||||
Db   62 LNIPNIIAAALTRNASAIHPGYGFLAENARFAEICADHHLTFIGPSPDSIRAMGDKSTAK 121

Qy 132 TAAKKAGLPVLAESTP-SKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRFVASPDELRKL 190
         | || ||||::|||:| ||||||| |||||||| |||||||||||||:||:||||:||
Db 122 ETMQRVGVPTIPGSDGLLTDVDSAAKVAAEIGYPVMIKATAGGGRGMRLVREPADLEKL 181

Qy 191 ATEASREAAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE 250
         || ||||||||| ||||||||:||||:||| |:|||||||| :|| ||||||||||||:
Db 182 FLAAQGEAEAAFGNPGLYLEKFIDRPRHVEFQILADAYGNVHLGERDCSIQRRHQKLLE 241

Qy 251 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 310
         || :| ||||| ||| ||| | ||:||| ||||||||||| ||| |||||||||||| |
Db 242 EAPSPALSADLRQKMGDAAVKYAQAIGYIGAGTVEFLVDATGNFYFMEMNTRIQVEHPVT 301

Db 359 YLPPGGPGVRDSHVTDYEIPPYYDSLIGKLIVWGATREEAIARMQRALRECAITGLPT 418

Qy 431 NIGFLRALLREEDFTSKRIATGFI 454
         :|:|::| :||| || ||||||::
Db 419 TLSFHQLMLQMPEFLRGELYINFV 442

RESULT 14
US-08-468-793-8
; Sequence 8, Application US/08468793
; Patent No. 6177267
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert
; APPLICANT: Gornicki, Piotr
; TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Arnold, White & Durkee
;   STREET: P.O. Box 4433
;   CITY: Houston
;   STATE: Texas
;   COUNTRY: United States of America
;   ZIP: 77210
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/468,793
;   FILING DATE: 06-JUN-1995
;   CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/422,560
;   FILING DATE: 14-APR-1995
;   APPLICATION NUMBER: US SN 07/956,700
;   FILING DATE: 02-OCT-1992
;   CLASSIFICATION: 800
;   APPLICATION NUMBER: PCT/US93/09340
;   FILING DATE: 30-SEP-1993
;   CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
;   NAME: Kitchell, Barbara S.
;   REGISTRATION NUMBER: 33,928
;   REFERENCE/DOCKET NUMBER: ARCD:152/KIT
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (512) 418-3000
;   TELEFAX: (713) 789-2679
;   TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 453 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-793-8

Query Match   17.4%; Score 1004.5; DB 3; Length 453;
Best Local Similarity 47.1%; Pred. No. 3.5e-71;
Matches 209; Conservative 72; Mismatches 158; Indels 5; Gaps 3;

Qy   12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 71
         ||:||||||||||||||||:|| |:: | |:|: | |||: |||||||||:||||||||
Db   3 FNKILIANRGEIALRILRTCEELGIGTIAVHSTVDRNALHVQLADEAVCIG-EAASSKSY 61

Qy   72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 131
         |:||||||||:|||||||||||||:|||:|||| | :||:||||:||||:||||||||||
Db   62 LNIPNIIAAALTRNASAIHPGYGFLAENARFAEICADHHLTFIGPSPDSIRAMGDKSTAK 121

Qy 132 TAAKKAGLPVLAESTP-SKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRFVASPDELRKL 190
         | || ||||::|||:| ||||||| |||||||| |||||||||||||:||:||||:||
Db 122 ETMQRVGVPTIPGSDGLLTDVDSAAKVAAEIGYPVMIKATAGGGRGMRLVREPADLEKL 181

```
Qy   311  EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 370
          |:|.:||.|:||:|::  |.|: |||::||| ||| :|||
Db   302  EMITGLDLIAEQIRIAQGEALR---FRQADIQLRGHAIECRINAEDPEYNFRPNPGRITG 358

Qy   371  YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 430
          |. ||||:||| :||  ||::|::|:   |:.||  ||||| ||::|
Db   359  YLPPGGPGVRVDSHVYTDYEIPPYDSLIGKLIVWGATREEAIARMQRALRECAITGLPT 418

Qy   431  NIGFLRALLREEDFTSKRIATGFI 454
          ::.:|:. :|   : | |:
Db   419  TLSFHQLMLQMPEFLRGELYTNFV 442
```

Search completed: September 24, 2003, 15:43:57
Job time : 21.8668 secs

```
                    GenCore version 5.1.6
             Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:        September 24, 2003, 15:41:22 ; Search time 36.2299 Seconds
                                               (without alignments)
                                               4760.992 Million cell updates/sec

Title:         US-09-974-973A-19
Perfect score: 5788
Sequence:      1 MSTHTSSTLPAFKKILVANR.........RVVVPAATKVEGGDLIVVVS 1140

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:      566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters:      566894

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries
```

```
Database :      Published_Applications_AA:*
         1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
         2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
         3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
         4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
         5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
         6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
         7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
         8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
         9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
        10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
        11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
        12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
        13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
        14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
        15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
        16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
        17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
        18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
Result              %      Query
No.   Score  Match  Length  DB  Length  ID                          Description
----  -----  -----  ------  --  ------  --------------------------  ------------------
  1   5788   100.0   1140   10   1140   US-09-974-973-19             Sequence 19, Appl
  2   5788   100.0   1140   10   1140   US-09-738-626-4265           Sequence 4265, Ap
  3   5788   100.0   1140   15   1157   US-10-045-072-2              Sequence 2, Appli
  4   5759    99.5   1157   10   1157   US-09-974-973-2              Sequence 2, Appli
  5   5759    99.5   1157   10   1157   US-09-974-973-4              Sequence 4, Appli
  6   2488    43.0   1147    9   1142   US-09-815-242-5468           Sequence 5468, Ap
  7   2472.5  42.7   1142    9   1142   US-09-815-242-10806          Sequence 10806, A
  8   2299    39.7   1073    9   1073   US-09-815-242-12361          Sequence 12361, A
  9   1036.5  17.9    447    9    447   US-09-767-479-6              Sequence 6, Appli
 10   1004.5  17.4    453    9    453   US-09-767-479-8              Sequence 8, Appli
 11    976.5  16.9    471    9    471   US-09-815-242-5215           Sequence 5215, Ap
 12    957    16.5    449    9    449   US-09-815-242-13885          Sequence 13885, A
 13    946    16.3    448    9    448   US-09-815-242-11160          Sequence 11160, A
 14    942.5  16.3    455    9    455   US-09-815-242-13617          Sequence 13617, A
 15    942    16.3   1171   15   1171   US-10-156-761-14226          Sequence 14226, A
 16    941.5  16.3    455    9    455   US-09-815-242-13364          Sequence 13364, A
 17    938    16.2    449    9    449   US-09-815-242-10330          Sequence 10330, A
 18    924    16.0    449    9    449   US-09-815-242-12063          Sequence 12063, A
 19    914    15.8    458    9    458   US-09-815-242-11321          Sequence 11321, A
 20    911    15.7    456    9    456   US-09-815-242-10924          Sequence 10924, A
 21    911    15.7    590   15    590   US-10-156-761-10874          Sequence 10874, A
 22    911    15.7    725   15    725   US-10-160-501-17             Sequence 17, Appl
 23    910    15.7    443    9    443   US-09-815-242-4963           Sequence 4963, Ap
 24    910    15.7    725   15    725   US-10-224-539A-2             Sequence 2, Appli
 25    910    15.7    725   15    725   US-10-224-539A-9             Sequence 9, Appli
 26    905    15.6    590   12    590   US-10-045-612A-25            Sequence 25, Appl
 27    897    15.5    455    9    455   US-09-815-242-11558          Sequence 11558, A
 28    892    15.4    590   12    590   US-10-045-612A-26            Sequence 26, Appl
 29    892    15.4    700   15    700   US-10-156-761-12811          Sequence 12811, A
 30    880.5  15.2    591   10    591   US-09-738-626-6940           Sequence 6940, Ap
 31    859.5  14.8    446    9    446   US-09-815-242-5418           Sequence 5418, Ap
 32    859.5  14.8    453    9    453   US-09-815-242-12562          Sequence 12562, A
 33    840    14.5    448    9    448   US-09-815-242-5806           Sequence 5806, Ap
 34    837    14.5    451    9    451   US-09-815-242-13127          Sequence 13127, A
 35    826.5  14.3    616   15    616   US-10-156-761-11400          Sequence 11400, A
 36    653    11.3    464   15    464   US-10-169-048-28             Sequence 28, Appl
 37    636    11.0    358    9    358   US-09-767-479-10             Sequence 10, Appl
 38    580.5  10.0   2257    9   2257   US-09-839-477-8              Sequence 8, Appli
 39    532.5   9.2   2257   12   2257   US-10-083-357-1328           Sequence 1328, Ap
 40    424.5   7.3    483   15    483   US-10-224-539A-5             Sequence 5, Appli
 41    421.5   7.3    262   15    262   US-09-815-242-5031           Sequence 5031, Ap
 42    380.5   6.6    158    9    158   US-09-895-913A-324           Sequence 324, App
 43    301.5   5.2    163   10    163   US-10-210-115-20             Sequence 20, Appl
 44    255     4.4   1073   16   1073   US-09-205-658-240            Sequence 240, App
 45    245     4.2    124    9    124                                Sequence 240, App
```

ALIGNMENTS

```
RESULT 1
US-09-974-973-19
; Sequence 19, Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
;   APPLICANT: Hanke, Paul D.
;   TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacter
;   FILE REFERENCE: 1533.1230001/MAC/RGM
;   CURRENT APPLICATION NUMBER: US/09/974,973
;   CURRENT FILING DATE: 2001-10-21
;   PRIOR APPLICATION NUMBER: US 60/239,913
;   PRIOR FILING DATE: 2000-10-13
;   NUMBER OF SEQ ID NOS: 19
;   SOFTWARE: PatentIn version 3.0
;   SEQ ID NO 19
;   LENGTH: 1140
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-974-973-19

  Query Match          100.0%;   Score 5788;    DB 10;   Length 1140;
  Best Local Similarity 100.0%;   Pred. No. 0;
  Matches 1140;   Conservative  0;   Mismatches  0;   Indels  0;   Gaps  0;

Qy      1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60

Qy     61 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120

Qy    121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180

Qy    181 VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280998
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4265
; LENGTH: 1140
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4265

Query Match   100.08%;   Score 5788;
Best Local Similarity   100.08%;   Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;   DB 10;   Length 1140;

```
Qy    1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60

Qy   61 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120

Qy  121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRF 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRF 180

Qy  181 VASPDELRKLATEASREAEAAFGDGAVYYERAVINPQHIEVQILGDHTGEVVHLYERDCS 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 VASPDELRKLATEASREAEAAFGDGAVYYERAVINPQHIEVQILGDHTGEVVHLYERDCS 240

Qy  241 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 300

Qy  301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360

Qy  361 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420

Qy  421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV 480

Qy  481 TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 540

Qy  541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600

Qy  601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 660

Qy  661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR 720

Qy  721 PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 780

Qy  781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQL 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQL 840

Qy  841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900

Qy  901 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA 960

Qy  961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLLRLPDVRTP 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLLRLPDVRTP 1020

Qy 1021 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080

Qy 1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATTASVDGKIDRVVVPAATKVEGGDLIVVVS 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATTASVDGKIDRVVVPAATKVEGGDLIVVVS 1140
```

RESULT 2
US-09-738-626-4265
; Sequence 4265, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:

; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO

Db　181　VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVHLYERDCS

Qy　241　LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN　300
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　241　LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN　300

Qy　301　PRIQVEHVTVEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG　360
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　301　PRIQVEHVTVEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG　360

Qy　361　FRPDTGTTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL　420
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　361　FRPDTGTTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL　420

Qy　421　AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV　480
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　421　AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV　480

Qy　481　TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT　540
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　481　TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT　540

Qy　541　FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE　600
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　541　FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE　600

Qy　601　LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA　660
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　601　LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA　660

Qy　661　IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR　720
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　661　IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR　720

Qy　721　PAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS　780
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　721　PAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS　780

Qy　781　LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL　840
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　781　LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL　840

Qy　841　SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF　900
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　841　SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF　900

Qy　901　AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA　960
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　901　AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA　960

Qy　961　DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLLRLPDVRTP　1020
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　961　DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLLRLPDVRTP　1020

Qy　1021　LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA　1080
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1021　LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA　1080

Qy　1081　GVVTVTVVAEGDEVKAGDAVAIIEAMKMEATTTASVDGKIDRVVVPAATKVEGGDLIVVS　1140
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1081　GVVTVTVVAEGDEVKAGDAVAIIEAMKMEATTTASVDGKIDRVVVPAATKVEGGDLIVVS　1140

RESULT 4
US-09-974-973-2
; Sequence 2, Application US/09974973
; Patent No. US2002017202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacter
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974,973
; CURRENT FILING DATE: 2001-10-21

Db　181　VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVHLYERDCS　240

Qy　241　LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN　300
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　241　LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN　300

Qy　301　PRIQVEHVTVEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG　360
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　301　PRIQVEHVTVEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG　360

Qy　361　FRPDTGTTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL　420
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　361　FRPDTGTTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL　420

Qy　421　AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV　480
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　421　AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV　480

Qy　481　TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT　540
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　481　TVNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT　540

Qy　541　FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE　600
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　541　FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE　600

Qy　601　LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA　660
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　601　LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA　660

Qy　661　IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR　720
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-10-045-072-2

RESULT 3
US-10-045-072-2
; Sequence 2, Application US/10045072
; Publication No. US20030027305A1
; GENERAL INFORMATION:
; APPLICANT: Sinskey, Anthony J.
; APPLICANT: Lessard, Philip A.
; APPLICANT: Willis, Laura B.
; TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
; FILE REFERENCE: 1533.0790002
; CURRENT APPLICATION NUMBER: US/10/045,072
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 09/677,575
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US 09/220,081
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1140
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-045-072-2

Query Match　　　　　100.0%;　　Score 5788;　　DB 15;　　Length 1140;
Best Local Similarity　100.0%;　　Pred. No. 0;
Matches 1140;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

Qy　1　MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSPHRSFASEAVR　60
　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1　MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSPHRSFASEAVR　60

Qy　61　IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV　120
　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　61　IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV　120

Qy　121　LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRF　180
　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　121　LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRF　180

Qy　181　VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVHLYERDCS　240
　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
; PRIOR APPLICATION NUMBER: US 60/239,913
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1157
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-974-973-2

Query Match          99.5%;   Score 5759;  DB 10;  Length 1157;
Best Local Similarity 99.4%;   Pred. No. 0;
Matches 1133;  Conservative 5;  Mismatches 2;  Indels 0;  Gaps 0;

Qy    1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   18 VSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 77

Qy   61 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   78 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 137

Qy  121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180
          ||||||||||||||||||||||||||||| :|||||||||||||||||||||||||||||
Db  138 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRF 197

Qy  181 VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240
          : |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db  198 VSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCS 257

Qy  241 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  258 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 317

Qy  301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  318 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 377

Qy  361 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  378 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 437

Qy  421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  438 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADV 497

Qy  481 TVNKPHGVRPKDVAAPIDKLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  498 TVNKPHGVRPKDVAAPIDKLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 557

Qy  541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  558 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 617

Qy  601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  618 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA 677

Qy  661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  678 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR 737

Qy  721 PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  738 PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 797

Qy  781 LSAIVAAFAHTRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  798 LSAIVAAFAHTRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 857

Qy  841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  858 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 917
```

```
Qy  901 AADPQKTDIPDSVIAFLRGELGNPPGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  918 AADPQKTDIPDSVIAFLRGELGNPPGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA 977

Qy  961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  978 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1037

Qy 1021 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1038 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1097

Qy 1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATTASVDGKIDRVVVPAATKVEGDLIVVVS 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1098 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATTASVDGKIERVVVPAATKVEGDLIVVVS 1157
```

```
RESULT 5
US-09-974-973-4
; Sequence 4, Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacter
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974,973
; CURRENT FILING DATE: 2001-10-21
; PRIOR APPLICATION NUMBER: US 60/239,913
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1157
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-974-973-4

Query Match          99.5%;   Score 5759;  DB 10;  Length 1157;
Best Local Similarity 99.4%;   Pred. No. 0;
Matches 1133;  Conservative 5;  Mismatches 2;  Indels 0;  Gaps 0;

Qy    1 MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   18 VSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 77

Qy   61 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   78 IGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 137

Qy  121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180
          |||||||||||||||||||||||||||||| :|||||||||||||||||||||||||||
Db  138 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRF 197

Qy  181 VASPDELRKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240
          :|||||||||||||||||||||||| |||||||||||||||||||| |||||||||||||
Db  198 VSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCS 257

Qy  241 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  258 LQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 317

Qy  301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  318 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 377

Qy  361 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  378 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 437

Qy  421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV 480
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db  438 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADV 497
```

```
Query Match             43.0%;  Score 2488;  DB 9;  Length 1147;
Best Local Similarity   45.2%;  Pred. No. 1.5e-190;
Matches  527;  Conservative 192;  Mismatches 384;  Indels  384;  Gaps  14;

Qy   13  KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL   72
             :|:.||  |||||.||  :|:||   : :|| |     :|:|:|   ::||.|   ||.:| :|
Db    4  KKLLVANRGEIAIRFRAAAELDISTVAIYSNEDKSSLHRYKADESYLVGSDLGPAESYL   63

Qy   73  DIDEIIGAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT   132
         :|: :| : ||:|.||||| |||||:|.|| || :|| :|||| : |  :| |||.:|||
Db   64  NIERIIDVAKQANVDAIHPGYGFLSENEQFARRCAEEGIKFIGPHLEHLDMFGDKVKART   123

Qy  133  AAKKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRFVASPDELRKLA   191
          | | :||||| |    || :|:|:| || : | |::||| | ::|||: | | |:  || |
Db  124  TAIKADLPVIPGTDGPIKSYELAKEFAEEAGFPLMIKAISGGGKGMRIVREESELEDAF   183

Qy  192  TEASREAEAAFGDGAVYYERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVEI   251
         : | |||| .||::.|| |||:|:||:|||||:::|:.||:|||:|||||.|||||||||||
Db  184  HRAKSEAEKSFGNSEVYIERYIDNPKHIEVQVIGDEHGNIVHLFERDCSVQRRHQKVVEV   243

Qy  252  APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLV--DEKGNHVFIEMNPRIQVEHIV   309
         || . | *| ||:|| || :: ::||  |||||||||   *|   |||.:||||:|||| :|
Db  244  APSVGLSPTLRQRICDAAIQLMENIKYYNAGTVEFLVSGDE---FFFIEVNPRVQVEHTI   300

Qy  310  TEEVTEVDLVKAQMRLAAGATL-KELGLTQDK-IKTHGAALQCRITTEDPNNGFRPDTG   366
         || || :|:||:|:  .|||.|  . |: | ||:  .| | : :||||||||||| . .|.||
Db  301  TEMVTGIDIVKTQILVRAGADLFGEEINMPQQKDITTLGYAIQCRITTEDPLNDFMPDTG   360

Qy  367  TITAYRSPGGAGVRLD-GAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTV   425
         |||||||.|| .|||| | : : | |:: :|||:|| : . :|:::| :|:. |:|  |:|
Db  361  TITAYRSSGGFGVRLDAGDGFQGAEISPYYDSLLVKLSTHAISFKQAEEKMVRSLREMRI   420

Qy  426  SGVATNIGFLRALLREEDFTSKRIATGFIADRPHLLQAPPADDEQGRILDYLADVTVNKP   485
         .||.|||.|| .:: ::.||| : :| :| :.|.|: | |. . |. |:.|| |.|.| 
Db  421  RGVKTNIPFLINVMKNKKFTSGDYTTKFIEETPELEFDIQPSLLDRGTKTLEYIGNVTIN--   478

Qy  486  HGVRPKDVAAPIDKLPNIKDLPLP---------------RGSRDRLKQLGPAAFARD   527
           ||:|:|:| : :  |:| | :  |:|                  .| | :| ||| |.||
Db  479  --------GFPNVEKRPKPDYELASIPTVSSSKIASFSGTKQLLDEVGPKGVAEW   525

Qy  528  LREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAM   587
         : |.||:|||||||||||||||||||||  :| | |:| ||:| | :|| | ||.:||:||
Db  526  VKKQDDVLLTDTTFRDAHQSLLATRVRVTKDMINIASKTADVFKDGFSLEMWCGATFPVAY   585

Qy  588  RFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRI   647
         .||:|:||:||: *|: .||||  ||||| | |||| ||| | : || : | :||:|||||
Db  586  NFLKENPWERLERLRKAIPNVLFQMLLRASNAVGYKNYPDNVIHKFVQESAKAGIDVFRI   645

Qy  648  FDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDP-NEKLYTLDYYLKMAEEIVKSG   706
         || || | :|: | :|: |  :| ::| | ||| ::| | ::|:||:|| |:|:|: :|
Db  646  FDSLNWVDQMKVANEAVQEAG-KISEGTICYTGDILNPERSNIYTLEYYVKLAKELEREG   704

Qy  707  AHILAIKDMAGLLRPAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVD   766
          |:|||||||||| | : :| |  :| :| ||:|:|||| | .||| : |:| :|.| |
Db  705  FHLAIKDMAGLLKPKAAYELIGELKAAVDLPIHLTHDTSGNGLLTYKQAIDAGVDLID   764

Qy  767  GASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGP   826
          . :.|||:||||| ::|: || .  | : :  | |::|  :|:. | |:|  :|.|:||
Db  765  TAVASMSGLTSQPSANSLYYALNGPFPRHLRTDIEGMESLSHYWSTVRTYYSDFESDIKSP   824
```

```
Qy  481  TVNKPHGVRPKDVAAPIDKLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT   540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  498  TVNKPHGVRPKDVAAPIDKLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT   557

Qy  541  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE   600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  558  FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE   617

Qy  601  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA   660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  618  LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPA   677

Qy  661  IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR   720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  678  IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLR   737

Qy  721  PAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS   780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  738  PAAVTKLVTALRREFDLPVHVTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS   797

Qy  781  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVIRHEIPGGQL   840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  798  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVIRHEIPGGQL   857

Qy  841  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF   900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  858  SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF   917

Qy  901  AADPQKYDIPDSVIAFLRGELGNPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA   960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  918  AADPQKYDIPDSVIAFLRGELGNPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDA   977

Qy  961  DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP   1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  978  DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP   1037

Qy  1021 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA   1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1038 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA   1097

Qy  1081 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS   1140
         |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db  1098 GVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVVPAATKVEGGDLIVVVS   1157
```

RESULT 6
US-09-815-242-5468
; Sequence 5468, Application US/09815242
; Patent No. US2002061569A1

```
Qy    73  DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT  132
              ::|||  |||||:||||||||  ||: ::  :|      ||: ||  |  |||  :||:
Db    62  DIENIIQIAKKSGADAIHPGYGFLSENLRFAERCEEEGIIFVGPKTHHLDIFGDKIKAKE  121

Qy   133  AAKKAGLPVLAEST-PSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKLA  191
           |||| ||  ||:   |:|:|:||||:|||:| : | ||||  |||||||||     ||||
Db   122  AAVAAGIASIPSDGPVATVEEVVAFGETHGFPIMIKAALGGGGRGMRVAHDAKEAARGY  181

Qy   192  TEASREAAAFGDGAVVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVEI  251
            ||||| || :||||   ||:|:|  |||||  ||:||||:| || | |||||||||||
Db   182  ERAKSEAKAAFGSDEVVEKYISNPKHIEVQILGDHGNVLHLFERDCSVQRRHQKVVEV  241

Qy   252  APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE  311
            |||||| |||| ||||: |    ::|  :|    ||| ||:| | :||||||:|||||::
Db   242  APCVSMNEEQRAAICSAAVQLMAHVGVVNAGTVEFLV-EGDQFYFIEVNPRVQVEHTITE  300

Qy   312  EVTEVDLVKAQMRLAAGATL-KELGL-TQDKIKTHGAALQCRITTEDPNNGFRPDTGTIT  369
            |  ||| | :: |||  ||  |:||   || | |  |||| ||||||||:|| | |  |:
Db   301  MITDIDIVISQLQIAQGLDLHKDMHLPKQNELTLKGAAIQCRITTEDPLNQFMPDTGKID  360

Qy   370  AYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGV  428
            ||||| | || || |     || | : || : ||: :|| :|  :: | ||||| |: ||
Db   361  TYRSPGGFGVRLDVGNAYSGYAVTPFYFDSLLVKVCTHGFSFEQAISKMQRCLKEFRIRGV  420

Qy   429  ATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVNKPHGV  488
            ||:|   |   :  ||||: | : | :| | ||||  || :|| |||||:||||:||| |
Db   421  KTNIPFLQNVVSYPAFQSGEAKTTFIDNTPELFEFPMRDRGNKTMKYIGEVTVNGFPGI  480

Qy   489  R--------------PKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAV  536
                          ||: |::||:|:|:||:|:|| |: :     |  |||  :|| || |
Db   481  ERTEKKYFEAPRVPTDIEVP-EKVITAKNI------------LDAQGATAVIDWKNQESVLM  530

Qy   537  TDTTFRDAHQSLLATRVRSFALKPAAEAVAKLT---PELLSVEAWGGATYDVAMRFLFE  592
            |||||||||||||||||||:||:||  ||| |||    |:|| ||||||||:|| |||
Db   531  TDTTFRDAHQSLLATRVRTQDFK----AIAGLTDVALPELFSSEMGGATFDVAYRFLTE  586

Qy   593  DPWDRLDELRREAMPNVNIQMLLRGRNTVGTYPYPDSVCRAFVKEAASSGVDIFRIFDALN  652
            ||| |::| | |: || ||::|||: :|   ||:|  :|   |:|||  ||| ||||||||
Db   587  DPWQRLRKIRQLMPNILLQMLFRGSNAVGYQNYPDNVIEEFIKESARQGVDVFRIFDSLN  646

Qy   653  DVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAI  712
            |||| || || | ::  :| |  |  :|:   ::| | | | ||  |: ::| |||||||
Db   647  WIPQMEKSIQVVRDTG-KIAEAAICYTGDINDPARAKYNVQYYLDMAKELENLGAQIIAI  705

Qy   713  KDMAGLLRPAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPL  772
            |||||||:|  ::|:  :|  |||::|||||||:|:|||||| :||| |||| || | |:
Db   706  KDMAGLLKPQAAYRLISELKAATDLPIHLHTHDTSGNGIITYSAATKAGVDIVDVAMSAM  765

Qy   773  SGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLLPFESCTPGPTGRVYR  832
            | || ||:|    :| | ||:|:|  || | :|:  | ||  ||| | |   ||| :| |
Db   766  SGATSQPSMNSLYYALVNGERTPTINIDNAQKINHYWEDVRMYYQPFENGLNAPQTEVYM  825

Qy   833  HEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVG  892
            ||:|||| |||||| |||| ||:|| || || |||||||||||| | |||||||| :|||
Db   826  HEMPGGYSNLQQQAKAVGLGHRWDEIKKMYHTVNLMFGDIVKVTPSSKVVGDMALFMVQ  885

Qy   893  AGVDPADFRAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPE  952
            : |||| ||  | ||   || |     :||||  ||  |  : ||  | |     | |  |
Db   886  NHLTEQDVVARGEELSFPESVVTFFQGDLGQPVGGFPKELQRIILKGR---PAFTERPG  941

Qy   953  EEQAHLDADDSKER----------RNSLNRLLFPKPTEEFLEHRRR---FGNTSALDDR  998
            |||:| | |||:           :|::| || ||||||||: :  :     :  ||| |
Db   942  DLAAPVDFAKVQEELAEKIGYQPKLEEVLSYLMYP----QVFLEYRQKYETFGDITLLDTP  998

Qy   999  EFFYGLVEGRETLLRLPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVE  1058
            | |:|| ||:| || ||||||||| ||||:|||||||||||||:||| :||||||||: |
Db   999  TFFNGIRQGETLEVQIERGKT-LIIRLDEIGEPDIDGNRVLFFNLNGQRREVLVKDASIK  1057

Qy  1059  SVTATAEKADSSNKGHVAAPFAG--VVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDG  1117
            |:|||||||:||:|||||||| |  ||:|:|||||||||||||| || ||| |::||||||
Db  1058  SAVQVKQKAEPTNKEQIGATMSSGSVLQVLVKRGDKVEKGQPLLITEAMKMETTIEARFAG  1117

Qy  1118  KIDRVVVPAATKVEGGDLIVVV  1139
            |:| |::|  |||||||||:::
Db  1120  FDGVIKQVTVNNGDTIATGDLLIEI  1144
```

```
Query Match      42.7%;  Score 2472.5;  DB 9;  Length 1142;
Best Local Similarity   45.8%;  Pred. No. 2.7e-189;
Matches  532;  Conservative 185;  Mismatches 386;  Indels  59;  Gaps  17;

Qy     13  KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSPASEAVRIGTEGSPVKAYL  72
               |:|||||||||||| ||||:|||:|:  |||: || |   | |||  ||| ||||||
Db      2  KKVLVANRGEIAIRIFRACTELDIRTVAIYAAEDEYSVHRFKADEAYLVGKEKKPIEAYL  61
```

```
Db  1118  TVDHIYVEEGEAISSGDLLLEV 1139
              :|:  |||  |  |||::| |

RESULT 8
US-09-815-242-12361
; Sequence 12361, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12361
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12361

Query Match             39.7%;  Score 2299;  DB 9;  Length 1073;
Best Local Similarity   44.8%;  Pred. No. 2.1e-175;
Matches  489;  Conservative  179;  Mismatches  361;  Indels  62;  Gaps  14;

Qy   87  DAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVL-AES 145
         |||::|||:||||| : |  : | ::|| |||  :|||:
Db    2  DAIHPGYGFLSENEQFARRCAEEGIKFIGPHLEHLDMFGDKVKARTTAIKADLPVIPGTD  61

Qy  146  TPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKLATEASREAEAAFGDG 205
          :| |||||  ::|| ::|:||||||||||||||| ::|||  :  |  :  |:||
Db   62  GPIKSYELAKEFAEEAGFPLMIKATSGGGGKGMRIVREESELEDAFHRAKSEAEKSFGNS 121

Qy  206  AVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRI 265
          :||| |: | |:|: :||||||  | : |: |||| | |||:|| |||
Db  122  EVYIERYIDNPKHIEVQVIGDEHGNIVHLFERDCSVQRRHQKVVEVAPSVGLSPTLRQRI 181

Qy  266  CADAVKFCRSIGYQGAGTVEFLV-DEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQM 323
         | || |  | ::|:| ||:||||  :: |: |: ||| |||| ::|:
Db  182  CDAAIQLMENIKYVNAGTVEFLVSGDE--FFFIEVNPRVQVEHTITEMVTGIDIVKTQI 238

Qy  324  RLAAGATL--KELGLTQDK-IKTHGAALQCRITTEDPNNGFRPDTGTITAYRSPGGAGVR 380
          :||||  : ::| ||  : | | ::|||:||||||||  | ||||||:|||:
Db  239  LVAAGADLFGEEINMPQQKDITLGYAIQCRITTEDPLNDFMPDTGTIIAYRSSGGPFGVR 298

Qy  381  LD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIGFLRALL 439
         || ||   ::|||:  | |||:  |:|  ::| |||  :|
Db  299  LDAGDGFQGAEISPYYDSLLVKLSTHAISFKQAEEKMVRSLREMRIRGVKTNIPFLINVM 358

Qy  440  REEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVNKPHGVRPKDVAAPIDK 499
          :||| |
Db  359                                              KNKFTSGDYTTKFIEETPELFDIQPSLDRGTKTLEYIGNVTIN------G 403

Qy  500  LPNIKDLPLP----------RGSRDRLKQLGPAAFARDLREQDALAVTDTTF 541
          :||: |  :||:  : | ::|  |  ::| | :: ::::|||:
Db  404  FPNVEKRPKPDYELASIPTVSSSKIASFSGTKQLLDEVGPKGVAEWVKKQDDVLLTDTTF 463

Qy  542  RDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDEL 601
         ||||||||||||| |: | ::|| | | |  |||||||||:||| ||||:||:||::|
Db  464  RDAHQSLLATRVRTKDMINIASKTADVFKDGFSLEMWGGATFDVAYNFLKENPWERLERL 523

Qy  602  REAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAI 661
         |:| |||  |||||  | |||  ||:|:  ||::|: | |||||||||:||::||||:| :
Db  524  RKAIPNVLFQMLLRASANVGYKNYPDNVIHKFVQESAKAGIDVFRIFDSLNWVDQMKVAN 583

Qy  662  DAVLETNTAVAEVAMAYSGDLSDP-NEKLYTLDYLKMAEEIVKSGAHILAIKDMAGLLR 720
         :|::| | |  : |:|:||| |:| | : ||||:||:|::| |: |:|||||||||||||:
Db  584  EAVQEAG-KISEGTICYTGDILNPERSNIYTLEYVKLAKELEREGFHILAIKDMAGLLK 642

Qy  721  PAAVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 780
         |:|| :|: ||:||:|||:|:|||||:|  |||  ||:||:|||: |||:|::|||||||
Db  643  PKAAYELIGELKAAVDLPIHLHTHDTSGNGLLTYKQAIDAGVDIIDTAVASMSGLTSQPS 702

Qy  781  LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
          ::: :|  |  |:|  :    |  :::|| :   :|::|:| |:: |:  : |: ||||:
Db  703  ANSLYYALNGFPRHLRTDIEGMESLSHYWSTVRTYYSDFESDIKSPNTEIYQHEMPGGQY 762

Qy  841  SNLRAQATALGLADREFLIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900
         |||  |:| :|||  | |  ::| :||: |  ::||||||||||:||:::| | :|:|
Db  763  SNLSQQAKSLGLGERFDEVKDMYRRVNFLFGDIVKVTPSSKVVGDMALYMVONDLDEQSV 822

Qy  901  AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDA 960
          :| ||:|  |||: ||:||:|:|  ||  |  || |  : :::: || |:|:  ::|:
Db  823  ITDGYKLDFPESVVSFFKGEIGQPVNGFNKDLQAVILKGQE---ALTARPGEYLEPVDF 878

Qy  961  DDSKE------------RRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRE 1009
         :||| :         |: :| :   : | :    :| ::  ||   |: | :: |:  :|
Db  879  EKVRELLEEEQQGPVIEQDIISYVLYPKVYEQYIQTRNQYGNLSLLDTPTFFGMRNGET 938

Qy 1010  TLIRLPDVRTPLLVRLDAISEPDDKGMRNVANVNGQIRPMRVRDRSVESVTATAEKADS 1069
          : :  | :   ::::| :|||| :|:|| :  |||  ::| : : ||:|| | |:|:|
Db  939  VEIEI-DKGKRLIIKLETISEPDENGNRTIYYAMNGQARRIYIKDENVHTNANVKPKADK 997

Qy 1070  SNKGHVAAPFAGVVT-VTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVYPAAT 1128
         || ||: | |:|||  |  : | :| | ||:||::|||||| |   :|| ||:  | |:|
Db  998  SNPSHIGAQMPGSVTEVKVSVGETVKANQPLIITEAMKMETTIQAPPDGVIKQVTVNNGD 1057

Qy 1129  KVEGGDLIVVV 1139
         ::  |||::
Db 1058  TIATGDLLIEI 1068

RESULT 9
US-09-767-479-6
; Sequence 6, Application US/09767479
; Patent No. US20010036654A1
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert
; APPLICANT: Gornicki, Piotr
; TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
;                     METHODS OF USE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
```

```
;                SOFTWARE: PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/09/767,479
;        FILING DATE: 22-Jan-2001
;        CLASSIFICATION: Unknown
;    PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: 08/468,793
;        FILING DATE: <Unknown>
;        APPLICATION NUMBER: US SN 07/956,700
;        FILING DATE: 02-OCT-1992
;        APPLICATION NUMBER: PCT/US93/09340
;        FILING DATE: 30-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;        NAME: Kitchell, Barbara S.
;        REGISTRATION NUMBER: 33,928
;        REFERENCE/DOCKET NUMBER: ARCD:152/KIT
;    TELECOMMUNICATION INFORMATION:
;        TELEPHONE: (512) 418-3000
;        TELEFAX: (713) 789-2679
;        TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 6:
;    SEQUENCE CHARACTERISTICS:
;        LENGTH: 447 amino acids
;        TYPE: amino acid
;        STRANDEDNESS: single
;        TOPOLOGY: linear
;    MOLECULE TYPE: protein
;    SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-767-479-6

Query Match          17.9%;  Score 1036.5; DB 9; Length 447;
Best Local Similarity    48.4%;  Pred. No. 1.3e-74;
Matches 216; Conservative  73; Mismatches 152; Indels  5; Gaps  3;

Qy    12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 71
Db     3 FDKILIANRGEIALRILRACEEMGIATIAVHSTVDRNALHVQLADEAVCIG-EPASAKSY 61

Qy    72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 131
Db    62 LNIPNIIAAALTRNASAIHPGYGFLSENAKFAEICADHHIAFIGPTPEAIRLMGDKSTAK 121

Qy   132 TAAKKAGLPVLAESTPSKNID-EIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKL 190
Db   122 ETMQKAGVPTVPGSEGLVETEQBGLELAKDIGYPVMIKATAGGGGRGMRLVRSPDEFVKL 181

Qy   191 ATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE 250
Db   182 FLAAQGEAGAAFGNAGVYIEKFIERPRHIEFQILADNYGNVIHLGERDCSIQRRNQKLLE 241

Qy   251 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 310
Db   242 EAPSPALDSDLREKMGQAAVKAAQFINTGAGTIEFLLDRSGQFYFMEMNTRIQVEHPVT 301

Qy   311 EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 370
Db   302 EMVTGVDLLVEQIRIAQGERLR--LTQDQVVLRGHAIECRINAEDPDHDFRPAPGRISG 358

Qy   371 YRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 430
Db   359 YLPPGGPGVRIDSHVYTDYQIPPYYDSLIGKLIVWGPDRATAINRMKKRALRECAITGLPT 418

Qy   431 NIGFLRALLREEDFTSKRIATGFIAD 456
Db   419 TIGFHQRIMENPQFLQGNVSTSFVQE 444
```

RESULT 10
US-09-767-479-8
; Sequence 8, Application US/09767479
; Patent No. US20010036654A1
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert

```
;                Gornicki, Piotr
; TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
;                    METHODS OF USE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Arnold, White & Durkee
;     STREET: P.O. Box 4433
;     CITY: Houston
;     STATE: Texas
;     COUNTRY: United States of America
;     ZIP: 77210
; COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/767,479
;     FILING DATE: 22-Jan-2001
;     CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/468,793
;     FILING DATE: <Unknown>
;     APPLICATION NUMBER: US SN 07/956,700
;     FILING DATE: 02-OCT-1992
;     APPLICATION NUMBER: PCT/US93/09340
;     FILING DATE: 30-SEP-1993
; ATTORNEY/AGENT INFORMATION:
;     NAME: Kitchell, Barbara S.
;     REGISTRATION NUMBER: 33,928
;     REFERENCE/DOCKET NUMBER: ARCD:152/KIT
; TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (512) 418-3000
;     TELEFAX: (713) 789-2679
;     TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 8:
;     SEQUENCE CHARACTERISTICS:
;         LENGTH: 453 amino acids
;         TYPE: amino acid
;         STRANDEDNESS: single
;         TOPOLOGY: linear
;     MOLECULE TYPE: protein
;     SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-767-479-8

Query Match          17.4%;  Score 1004.5; DB 9; Length 453;
Best Local Similarity    47.1%;  Pred. No. 5.1e-72;
Matches 209; Conservative  72; Mismatches 158; Indels  5; Gaps  3;

Qy    12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 71
Db     3 FNKILIANRGEIALRLRTCEELGIGTIAVHSTVDRNALHVQLADEAVCIG-EAASSKSY 61

Qy    72 LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV 131
Db    62 LNIPNIIAAALTRNASAIHPGYGFLAENARFAEICADHHLTFIGPSPDSIRAMGDKSTAK 121

Qy   132 TAAKKAGLPVLAESTP-SKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKL 190
Db   122 ETMQRVGVPTIPGSDGLLTDVDSAAKVAAEIGYPVMIKATAGGGGRGMRLVREPADLEKL 181

Qy   191 ATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE 250
Db   182 FLAAQGEAEAAFGNPGLYLEKFIDRPRHVEFQILADAYGNVVHLGERDCSIQRRHQKLLE 241

Qy   251 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVT 310
Db   242 EAPSPALSADLRQKMGDAAVKVAQAIGYIGAGTVEFLVDATGNFYFMEMNTRIQVEHPVT 301

Qy   311 EEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA 370
Db   302 EMITGLDLIAEQIRIAQGEALR---FRQADIQLRGHAIECRINAEDPEYNFNPGRITG 358
```

```
Qy  308 TVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGT 367
        |:||:.:||::|::| :|..|   || ||| |:|
Db  296 TITEEIFTGIDVVREQIRIASGL---ELSVKQDDIVHRGYALQFRINAEDPKNNFLPSFGK 352

Qy  368 ITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSG 427
        |:| :.:||| |.||  :|| |  |||||||.:||| |..| ||||.|| | |:|
Db  353 ITRYYAPGGGVGVRTDIAITYGYTIPPYYDSMCLKLIVWALTWEEALDRGLRALDMRVQG 412

Qy  428 VATNIGFLRALLREEDFTSKRIATGFIADHPHLLQ 462
        | |..|   |:| :..| ..| |:::
Db  413 VKTTAPYTQEILRNPEFRSGQFNTSFVESHPELTQ 447
```

RESULT 12
US-09-815-242-13885
; Sequence 13885, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13885
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-13885

Query Match   16.5%; Score 957; DB 9; Length 449;
Best Local Similarity 47.6%; Pred. No. 3.3e-68;
Matches 206; Conservative 61; Mismatches 158; Indels 8; Gaps 5;

Qy   14 KILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLD 73
         ||:|||||||| |..|||.||:.||| : ..||| ||..:|| | |.|::.| :||:
Db   4 KIVIANRGEIAILRACKELGIKTVAVHSSADRDLKHVLLADETVCIGPAPS--VKSYLN 62

Qy   74 IDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVTA 133
         | :|:||| .|:|.|:.|||||||||| |.| |.:||:|||| ::: | |||| ::||
Db   63 IPAIISAAEITGAVAIHPGYGFLSENAFAQVERSGFIFIGPKADTIRLMGDKVSAITA 122

Qy 134 AKKAGLPVLAESTPSKNIDEIVKSAEGQT--YPIFVKAVAGGGGRGMRFVASPDELRKLA 191
         .|||| .| .. . ::|: .:: : ... |||::|| | ||||||| | . .||.::
Db 123 MKKAGVPTVPGSDGPLGDDMNANRAHAKRIGYPVIIKASGSGGGRGMRVVRSDAELAQSI 182

Qy 192 TEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVEI 251
         : :.||| ||.: |:.:: :.:||:|||:|:| | |.|::|.||||::||||||||:
Db 183 SMTKAEAKAAFSNDMVYMEKYLENPRHIEIQVLADGQGNAIYLAERDCSMQRRHQKVVEE 242
```

```
Qy  371 YRSPGGAGVRLDGAAQLGGETAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT 430
        |:||:.:||||||||||||||||||:: :|..| ||||.|| | |:|
Db  359 YLPPGGPGVRVDSHVYTDYEIPPYYDSLIGKLIVWGATREEAIARMQRALRECAITGLPT 418

Qy  431 NIGFLRALLREEDFTSKRIATGFI 454
        .:.|   |:| :..| ..| |:::
Db  419 TLSFHQLMLQMPEFLRGELYINFV 442
```

RESULT 11
US-09-815-242-5215
; Sequence 5215, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/205,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5215
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5215

Query Match   16.9%; Score 976.5; DB 9; Length 471;
Best Local Similarity 45.7%; Pred. No. 9.7e-70;
Matches 208; Conservative 73; Mismatches 159; Indels 15; Gaps 5;

Qy   13 KKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL 72
         ||||:||||||||| |..|| || :||| :| .| | ::.|.:|| : | |:|| | |
Db   3 KKILIANRGEIAVRIVRACAEMGIRSVAVYSEADRHALHVKRADEAHSIGAD--PLAGYL 60

Qy   73 DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT 132
         : .:::.|| .:.:||:.|||||||||| || :|.:.||||||: .|: |.|| ..
Db   61 NPRALVNLAVESGCDALHPGYGFLSENAELAEICAERGIKFIGPSAQVIRRMGDKTEARR 120

Qy 133 AAKKAGLPVLAESTPS----KNIDEIVKSAEGQTYPIFVKAVAGGGRGMRFVASPDEL 187
         :.| |.: ::   ..   :: ::.:||: .|||:.|||::|||||||| ...||:|
Db 121 SMIAAGVP----CTPGTEGNVADLAEALREAERIGYPVMLKATSGGGGRGIRRCNSREEL 176

Qy 188 RKLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVHLYERDCSLQRRHQK 247
         .:||: :.|| ||.|| : |:.:: |:||:|||.|:|.| .|.|||:||||:|||| ||
Db 177 EQAFPRVISEAIKAFGSAEVFLEKCIVNPKHIEAQILADSFGNTVHLFERDCSIQRRNQK 236

Qy 248 VVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQYEH 307
         ::|||.: .| .| ||.| .:||: : |:|:: |||||||: :.| | :|||| |:| |||
Db 237 LIEIAPSPQLTPEQRAYIGDLAVRAAKAVGYENAGTVEFLLAD-GEVYFMEMNIRVQVEH 295
```

```
Qy  252  APAQHLDPELRDRICADAVKFCRSIGYQQAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTE  311
         ||| ||| ||:|| | :||||| |||   |||| |||| : |:||||||| ||||| |||
Db  243  APAPGITPELRRYIGERCAKACVDIGYRGAGTFEFLF-ENGEFYFIEMNTRIQVEHPVTE  301

Qy  312  EVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAY  371
         :|:| | |||||||||||||||||||  :|||||: |::|:|||||||| |:||||| :
Db  302  MITGVDLIKEQLRIAAG--QPLSITQDEVVVRGHAVECRINAEDPNT-FLPSPGKITRF   357

Qy  372  RSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATN  431
         ::||| ||||| ||||||||||||||:| :|  |:|: | :|::|: ||:|| | :| |
Db  358  HAPGGFGVRWESHIYAGYTVPPYYDSMIGKLICYGENRDVAIARMKNALQELIIDGIKTN  417

Qy  432  IGFLRALLREEDF  444
         | ||| : ::||
Db  418  IDLQTRIMNDEHF  430
```

```
Query Match             16.3%;  Score 946;  DB 9;  Length 448;
Best Local Similarity   47.6%;  Pred. No. 2.5e-67;
Matches  208;  Conservative  61;  Mismatches  154;  Indels  14;  Gaps  7;

Qy   13  KKIIVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL   72
         :|::||||||||:|| | |:::|||| || |||   || |  |:||||||::|| ::||
Db    3  EKVVIANRGEIALRILRACKELGIKTVAVHSTADRDLKHVLLADETICIGPAPS-AKSYL   61

Qy   73  DIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLIGDKSRAVT  132
         :|  || || ||||||:|||||||||||  :| | || |||||: :|:||:||||||:| :
Db   62  NIPAIIAAEVTGADAIHPGYGFLSENADFAEQVERSGFTFIGPTADVIRLMGDKVSAIK   121

Qy  133  AAKKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKL   190
         | |||| | | :  |  | ||:  |||:|:|  |::||||||||||||||||:||| | ::
Db  122  AMKKAGVPCVPGSDGPVSNDIAKNKEIAKRIGYPIIIKASGGGGRGMRVVRSEDALEES  181
```

```
Qy  191  ATEASRBAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE  250
         |||  |:|||||:||:| |: :::||:||||:|| |:||:||||||||||:||||||||||
Db  182  IAMTKAEAKAAFNNDMVYMEKYLENPRHVEIQVLADTHGNAVYLAERDCSMQRRHQKVVE  241

Qy  251  IAPAQHLDPELRDRI---CADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEH  307
         :|||  | | :|||    |:  | | |||| ||| |||||| | :||:|||||| ||||||
Db  242  EAPAPGITEEVRRDIGSRCANA--CVEIGYRGAGTFEFLY-ENGEFYFIEMNTRIQVEH   297

Qy  308  TVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGT  367
         ||| | |:|  || ||||||||||||| |:||:||||| |::|:|| |||| |:|||||||
Db  298  PVTEMITGVDLVKEQLRIAAGL--PISFKQEDIKVKGHAMECRINAEDPKT-FLPSPGK   353

Qy  368  ITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSG  427
         :  | ||| ||||| ||||||||||||||::| :||  ||: | :|:::||:||:|| : |
Db  354  VNHLHSPGGLGVRWDSHVYGGYTVPPHYDSMIAKLITYGDTREVAIRRMQNALSETIIDG  413

Qy  428  VATNIGFLRALLREEDF  444
           || | ||| : ::||
Db  414  IKTNIPLHELILIEDENF  430
```

```
Query Match             16.3%;  Score 942.5;  DB 9;  Length 455;
Best Local Similarity   46.0%;  Pred. No. 5e-67;
Matches  206;  Conservative  75;  Mismatches  160;  Indels  7;  Gaps  5;

Qy   12  FKKIIVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY   71
         | ||:||||||||||:: || ||||:||||:| || ||| |  |:| :||   |:  ::|
Db    2  FRKILIANRGEIAVRIIRARELGIATVAVYSTADKEALHTLLADEAVCIG-PGKATESY   60

Qy   72  LDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAV  131
         | ||  |:||   ||||:|||:|||||:   |  |  |||:|||| | |:||  |||:| |
Db   61  LNINAVLSAAVLTEAEAIHPGFGFLSENSKFATMCEEVGIKFIGPSGHVMDMMCGDKINAR  120
```

```
Qy  132 TAAKKAGLPVLAESTPS-KNIDEIVKSAEGQTYPIFVKAVAGGGRGMRFVASPDELRKL 190
        |||  ::|  ::| || |  :||  ||||:|||  | :| ||||| |:|  |  |:|:|
Db  121 AQMIKAGVPVIPGSDGEVHNSEEALIVAEKIGYPVMLKASAGGGKGIRKVEKPDDLVSA 180

Qy  191 ATEASREAAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE 250
        ||:||  |  |:| |:| :||  |  :|||||||| :| ||||| ||||||:::|||:|
Db  181 FETASSRPAKANYCNGAMYIERVIYPARHIEVQILGDEHGHVIHLGERDCSLQRNNQKVLE 240

Qy  251 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDE-KGNHVFIEMNPRIQVEHTV 309
        ::|: |:  || || |||||: :|:|||:||||||||:|| :: |::::||| |:|||:|
Db  241 ESPSIAIGKTLRHEIGAAAVRAAEFVGYENAGTIEFLLDEASSNFYFMEMNTRVQVEHPV 300

Qy  310 TEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTIT 369
        ||  : ||:|| |::  |   : |:  | ||:   | | :||| ||:| ||  |:|| ||
Db  301 TEFVSGVDIVKEQICIRAG--QPLSVKQEDIVLRGHAIECRINAENPAFNFAPSPGKIT 357

Qy  370 -AYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGV 428
         | | ||| |:|||:| | |  |: ||||:: ||  :| :||  ::||||| :| : ||
Db  358 NLYLPSGGVGLRVDSAVYPGYTIPPYYDSMIAKIIVHGENRFDALMKMQRALYELEIEGV 417

Qy  429 ATNIGFLRALLREEDFTSKRIATGFIAD 456
        :|| :| | |: :|: |   ||| : :|
Db  418 QTNADFQLDLISDRNVIAGDYDTSFLME 445


RESULT 15
US-10-156-761-14226
; Sequence 14226, Application US/10156761.
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT:   OMURA, SATOSHI
; APPLICANT:   IKEDA, HARUO
; APPLICANT:   ISHIKAWA, JUN
; APPLICANT:   HORIKAWA, HIROSHI
; APPLICANT:   SHIBA, TADAYOSHI
; APPLICANT:   SAKAKI, YOSHIYUKI
; APPLICANT:   HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14226
; LENGTH: 1171
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14226

Query Match          16.3%;    Score 942;  DB 15;  Length 1171;
Best Local Similarity 28.1%;   Pred. No. 2.6e-66;
Matches   368;   Conservative 145;   Mismatches 471;   Indels 326;   Gaps 44;

Qy   12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVK-A 70
        | :||||||||||||  | ||| ||| |||| :  |||   |  :| ||| |  ||:  |
Db    3 FDTLLVANRGEIAVRVIRTARELGLRTVAVYSDPDRSAPHVRLADEAVRLGP--APAKES 60

Qy   71 YLDIDEIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRA 130
        ||| : ::  || | || | ||||||| :|  ||:| : | | ||||| || : | | |
Db   61 YLDADLVLKAAKDTGAGAIHPGYGFLSEDAAFARRCEDAGIVFVGPTPEQLELFGAKHTA 120

Qy  131 VTAAKKAGLPVLAEDSTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRFVASPDELRKL 190
        :: |  ||:|:|  ||  |  |::|:   :   | | ::| ||||: |  :| ::| |:|
Db  121 RAAAEAAGVPLAPGTGLLASLDEALDAASRIGYPVMLKATGGGGGIGMSACRSAAELAES 180

Qy  191 ATEASREAAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQKVVE 250
        |   : | ::| |:|:|:||  : ::|:||||:|| | ||:|| :|||||||::|||:|
Db  181 WERVRRVTAASFSSAGIFLERLVEHARHVEVQVFGDGRGRVVLGDRDCSLQRRNQKVVE 240

Qy  251 IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVD-EKGNHVFIEMNPRIQVEHTV 309
```

```
Qy  241 EAPAPGLPSHVRDHLASAARDLCAAVGYRSAGTVEFVYDAAREEAVFLEVNTRLQVEHPV 300

Qy  310 TEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTIT 369
        ||| : ||||  ::||   ::  |   | :||:  | |      | ||| | ||| |: |
Db  301 TEEIYGVDLVAWMLRLAGG---DADVVRDPGPPRGHAVRARVYAEDPSREHRPSAGLLT 356

Qy  370 AYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVA 429
        : | | |: | ||:  : | |:|: : |:| |  : |:|  | ::| :  : : :  |:|
Db  357 RVEFP--QGVRVDGWVETGTEVTTSYDPMLAKVVAYGSDRAHALARLDEALARTRVDGIE 414

Qy  430 TNIGFLRALLREEDFTSKRIATGFIA------DHP------HL 460
        ||:| :|| | : | :   :|:  :|      | |      ||
Db  415 TNLGLVRAALADHGLRTATHTTATLAKITDPTPRVEVVAAGTLTTVQDWPGRTGYWQVGV 474

Qy  461 LQAPPADD-------------------------------EQGR 487
        |:| |  |                               | ||
Db  475 PPCGPMDDRSFRLGNRALGNEEGAPGLECTLRGPALRFTHATTVCVTGAPAPVTVDGTPV 534

Qy  488 VRPKDVAAPIDKL------------PNIKDLPLPRGSRDRLKQLGPAAF-----ARDLR 529
        : | :  || |:|            |:|  | | | :::   |||  |       ||:|
Db  535 AQWEPVTVPAGALLEVGAPAEPGLRTYVLFAGGLDVPAFLGSAATFTLGRFGGHGGRALR 594

Qy  530 EQDALAVDTTFRDA-----------HQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYD 584
        : | | ||: :||             : :|  | | |||| ||:||| |||:|||| |:  | ||:
Db  595 TGDVLHGGSVTDRGAPVPAGDRPVFATEWHVGALEGPHAAPEFTED--DIRDFYAAGWK 652

Qy  585 V----AMRFLFEDP-WDRLDELREAMPNVNIQMLLRGRNTVGYT--PY--PDS--628
        :     : :|:: |  |:  :      :: :|  | | |  ||    |   |
Db  653 VIHFNSARTGVRLVGPKPRWARSDGGEAGLHPSNIHDTPYSVGAVDYTGDMPVLLGPDGPS 712

Qy  629 ----VCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSD 684
            | | || :|||| || || |||:| || | | :| | : :| |:| |  |||||
Db  713 LGGFVCPAITVATAERW------KLGQLRPG-DTVRFT------ 742

Qy  685 PNEKLYTLDYLKMAEEIVKSGAHILAIKDMAGLLRPAAVT------KLVTALRREFD 736
                     ::|   : || ||:|  |||||:|||: |       :|| | | ||
Db  743 -------------PLAEDGSTRPAIVDGGVLARDGDVTYRSGDDNLLVEFGPMQLD 786

Qy  737 LPVHVTHDTAGGQLATYFAAAQAGADAV-------DGASAPLSGTTSQPSLS 782
        |  : : |::| | |   :   ::|: :|       | | |:  :   | :|
Db  787 LALRMRVH--------ALMEAVAEAGLDGVTDLTPGIRSLQIRTDPRRLPLR----ELL 833

Qy  783 AIVAAFAHTRRDTGLSLEAVSDLEPYWEAV--RGLYLPFESGTPGPTGRVYRHEIPGGQ 839
        |:| ||    ||| |:| : ::  |   |:  |||  | | |  ||| |:  :|||||
Db  834 AVVR------ETVSALPPTDQLVVPSRTVHLPLSWDDPATREAIERY------ 874

Qy  840 LSNLRAQA------TALGLADRFELIEDNYAAVNE---MLGRPTKVTPSSKVVGDIAL 888
        :: |  |:      :| |  |::|: |:|  :| :    :||| ||| ||| |||: |
Db  875 MAGVRDDAPWCPWNIEFIRRVNGLESVADVDTVFDAEYLVLG------LGDVYL 923

Qy  889 HLVGAGVDPADFAADPQ-----KYDIPDSVIAFLRGELGNP-------PGGWPEPLR 933
        ::  |||: |   | |      ||:|  ::  : ::  |::         |||  :||
Db  924 ---GA---PVATPLDPRHRLVTTKYNPARTWTAENSVGIGGAYLCVYGMEGPGGYQFVGR 977

Qy  934 TRALEGRSEGKAPLTEVPEEQAHLDADDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTS 993
        |   || | ::    : |  :  |::|                ||: ||   |::|
Db  978 T----------TQVMSGWQ------QRGAF-----EPGSPWL--LRF----1001

Qy  994 ALDDREFFYGLVEGRETLIRL-PDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRV 1052
         :|   | : |  |:: |:   ||: | :|:|| | ::   : |   : : :| : | |
Db 1002 --FDRIKWYAV--GPDELLDLRADIASGRFVPRVEEGEFSLAGYESFLAANAGSIAAFRA 1057

Qy 1053 R-------DRSVESVTATAEKADSSNKGHVAAPFAGVVT---VTVAEG 1090
        |        | :  ||| ||: | ::: | |  | ||||    | |
Db 1058 RQGAAFAAERDAWERAGEFARAEAEATAPEAPPAEVTVPVGGALVEAEFAASVWQLNVPG 1117

Qy 1091 DEVKAGDAVAIIEAMKMEAITTASVDGKIDRVVPAATKVEGGDLIVVVS 1140
        | |:||| | :||||||: : || :| |||| | | |||  |::||:
Db 1118 DAVTAGQPLLTLEAMKMESRVPAPMAGVVTEILARPGDQVAAGTALVVLA 1167
```

Search completed: September 24, 2003, 15:53:13

Job time : 42.2299 secs